

138758

STIC-Biotech/ChemLib

From: Chan, Christina
Sent: Friday, November 26, 2004 3:51 PM
To: Schnizer, Richard; STIC-Biotech/ChemLib
Subject: RE: 10/798,532 RUSH

Please rush. Thanks Chris

Chris Chan

TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841
Remsen, 3E89

-----Original Message-----

From: Schnizer, Richard
Sent: Friday, November 26, 2004 7:14 AM
To: Chan, Christina
Subject: 10/798,532 RUSH

Please authorize the following RUSH search request:

Please search the commercial and interference databases for SEQ ID NOS: 1-4 from 10/798,532. Please also search for nucleic acids that could encode the polypeptides of SEQ ID NOS: 2 and 4.

Thank you-

Richard Schnizer, Ph.D.
Patent Examiner
Art Unit 1635
Remsen 2D30
571-272-0762
Mail Box 2C18

STAFF USE ONLY

Searcher: Noble
Searcher Phone: 2-
Date Searcher Picked up: Nov 30 2004
Date Completed: 3
Searcher Prep/Rev. Time: 3
Online Time: 3

Type of Search
NA Sequence: # 4
AA Sequence: # 2
Structure: #
Bibliographic:
Litigation:
Patent Family:
Other:

Vendors and cost where applicable

STN:
DIALOG:
QUESTEL/ORBIT:
LEXIS/NEXIS:
SEQUENCE SYSTEM: COMPUGER
WWW/Internet:
Other(Specify):

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2004, 11:23:30 ; Search time 5310 Seconds

(without alignments)
8646.716 Million cell updates/sec

Title: US-10-798-532-1

Perfect score: 1260

Sequence: 1 atcgaggagatccgcgtcgcgc.....tgatcaaaccttgcttaa 1260

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 1821965908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_esc1:*
2: gb_esc2:*
3: gb_hic:*
4: gb_esc3:*
5: gb_esc4:*
6: gb_esc5:*
7: gb_esc6:*
8: gb_gsel:*
9: gb_gsel2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1115	88.5	1763	3	AK013347	Mus muscu
2	886.6	70.4	2846	3	AK030152	Mus muscu
3	861.8	68.4	1079	4	BM468107	AGENCOURT
4	723.8	57.4	917	4	BM451184	AGENCOURT
5	624.4	49.6	1015	6	BY712344	BY712344
6	619.2	49.1	657	4	BG547675	602575520
7	617.8	49.0	899	2	BE895119	601436041
8	616.4	48.9	772	4	BI562420	603256316
9	610.4	48.4	932	5	B0541306	AGENCOURT
10	595	47.2	875	4	BI259568	602968330
11	593.2	47.1	834	7	CF593363	AGENCOURT
12	585.2	46.4	682	4	BI333581	602997065
13	583.8	46.3	906	4	BG386905	602454704
14	579.6	46.0	888	5	BQ225071	AGENCOURT
15	560	44.4	602	7	CO737403	S1LT02c17
16	558	44.3	876	5	BU184346	AGENCOURT
17	557.6	44.3	779	5	BU447831	603767571
18	550.8	43.7	844	4	BG476616	602517423
19	519.8	41.3	824	2	BE568528	601342425
20	516.8	41.0	823	7	CO771206	testis ES
21	516	41.0	781	7	CN232453	WB074D11
22	499.4	39.6	670	7	CN359673	170004245
23	498.6	39.6	717	5	BU385761	603582650
24	496.8	39.4	523	2	BF929410	IL2-NT020

25	495.2	39.3	896	7	CO773091 testis ES
26	494.6	39.3	928	5	BO424438 AGENCOURT
27	494.2	39.2	1003	5	BO922020 AGENCOURT
28	482.4	38.3	716	7	CF745774 UI-M-GV0-
29	471.8	37.4	755	5	BU371858 603568233
30	469.8	37.3	794	5	BU205357 603104569
31	468.8	37.2	785	6	CB990192 AGENCOURT
32	468.2	37.2	824	4	BI253163 602973284
33	465	36.9	899	5	BU707284 UI-M-FR0-
34	463.6	36.8	870	5	BQ216813 AGENCOURT
35	463.4	36.8	884	5	BX281211 EX281211
36	460	36.5	696	5	BU439278 604147528
37	456.6	36.2	483	2	BF929411 IL2-NT020
38	454.2	36.0	631	5	BU112717 603131421
39	451.8	35.9	743	4	BI562440 603256324
40	450.6	35.8	933	2	BE886040 601507192
41	450.4	35.7	986	5	BU751496 CH4#002_E
42	433.2	34.4	820	7	CK597557 AGENCOURT
43	427.4	33.9	672	5	BX261106 BX261106
44	420	33.3	604	5	BX261107 BX261107
45	413.2	32.8	570	5	BU293329 604166812

ALIGNMENTS

RESULT 1	AK013347	1763 bp	mRNA	linear	HTC 03-APR-2004
LOCUS	AK013347				
DEFINITION	Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length enriched library, clone:281045404 product:kinase interacting with leukemia-associated gene (statmin), full insert sequence.				
ACCESSION	AK013347	GI:12850651			
VERSION	AK013347.1				
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Mus.				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subcloning of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	1 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subcloning of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)				
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitahara, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.				
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer				
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)				
MEDLINE	20530913				
PUBMED	11076861				
AUTHORS	4 THE RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.				
REFERENCE	Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)				
TITLE	5 THE FANTOM Consortium and the RIKEN Genome Exploration Research				
JOURNAL					
REFERENCE					
AUTHORS					

JOURNAL REFERENCE	TITLE
ADACHI, J., ALZAWA, K., AKAHIRA, S., AKIMURA, T., ARAI, A., AONO, H., ARKAWA, T., BONO, H., CARINCI, P., FUKUDA, S., FUKUNISHI, Y., FUNURO, M., HANAGAKI, T., HARA, A., HAYASHI, N., HIROUCHI, K., HITACKA, T., HORI, F., IMOTANI, K., ISHII, Y., ITOH, M., IZAWA, M., KANEKAWA, T., KATO, H., KAWAI, J., KOJIMA, Y., KOMO, H., KONDA, M., KOYA, S., KURIHARA, C., MATSUMURA, T., MIYAZAKI, A., NISHI, K., NOMURA, K., NUMATAKI, R., OHNO, M., OKAZAKI, Y., OKIDO, T., OWA, C., SATO, H., SATO, R., SAKAI, C., SAKAI, K., SANO, H., SASAKI, D., SHIBATA, K., SHIBATA, Y., SHINAGAWA, A., SHIRAKI, T., SOGABE, Y., SUZUKI, H., TAGAMI, Y., TAGAWA, A., TAKAHASHI, F., TANAKA, T., TEIJIMA, Y., TOYA, T., YAMAMURA, T., YEANISHI, A., YOSHIDA, K., YOSHINO, M., YURAMATSU, M. and HAYASHIZAKI, Y.	Group Phase I & II Team. Analysis of the mouse transcripome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002) 6 (bases 1 to 1763)
COMMENT	Direct Submission Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
TITLE	Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGGACATCTCCAGTTAAATTAAATCATCCCCCCC3'] cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 7.5 and subtraction to Rot = 37.5. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGGACATCTCCAGTTAAATTAAATCATCCCCCCC3']. cDNA was cleaved with XhoI and SacI. Cloning hosts, 5' end: XhoI; 3' end: SacI. Host: SOLR.
JOURNAL	Location/Qualifiers 1. 1763 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM DB:2810454004" /db_xref="taxon:10090" /clone="2810454004" /issue_type="whole body" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="10, 11 days embryo"
CDS	/note="unnamed protein product; kinase interacting with leukemia-associated gene (statmin) (MGI:MGI:1341508) putative" /codon_start=1 /protein_id="BAB28802.1" /db_xref="GI:12850652" /translation="MAGSGCGMGAEPRFLFAFGRLWVOVSRLSGSSASVRYRCGGTTPGSPGALKOPLPGTTGAAGAAYGFRKERALOLQGHRIIVLYIGFTIFESP NVPERCLLDELTVSELVSLYSHGCSMMNIOGARDEALAPFHSHGVYADKL PRINTMSAENECEFLIDRGLSFKNGDVKYITODGYRAPAEALAOAGOSDFE CTSAVDMSLGITILTFEMSGSKLTGTTRSDGRKANSAIIDHTASAVYNNAIPATH LRDIKSKMLDDPGRRI PAELMALCSPPFSITPAFIHDVLWLPVPVRLNLVVDDLV ENDEYEVEDVEDEKQCKGPVNVSLVPEKNPGRGVFEVYANAAGSQAOKXLTGR MFDPKFVAIFYPIPSAYKRGLYXTLL"
ORIGIN	Query Match 88.5%; Score 1115; DB 3; Length 1763; Best Local Similarity 92.8%; Pred. No. 0; Matches 1169; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

OY	1	ATGGGCGGGATATCGGCTGGCGCTGGGGGCGCGGAGCGCGCGCTTTTCTGAGAGCCCTTGGCGG	60
Db	159	ATGGGCGGGGCTCGGCTGGCGCTGGGGGCGCGGAGCGCGCGCTTCTTGAGAGCCCTTGGCGG	218
OY	61	CGGCTGTGGCAGGTACAGAGCGCTGTGGGTAGCGGCTCTCCGCTCGGTATCGGGTT	120
Db	219	CGGCTGTGGCAGGTACAGAGCGCGCTGGGGCAGCGGCTGTGGGCTCGGTATCACGGGTG	278
OY	121	CGGCTGTGGGCGCAACCCCTGGGCTGGCCCCCGGCGCCCTCAGCAGTTCTTGGCCGCAAGA	180
Db	279	CGGCTGTGGGTTACCCCGGGCTCGCCCCCGGCGCCCTCAAGACAGTTCTCTCCTCGGGA	338
OY	181	ACCAACGGGGCTGGCGGCTCTGCGCGCGAGTATGGTTTCCGCAAGAAGAGGGCGGCGCTG	240
Db	339	ACCAACGGGGGCTGGGGGCTCGGCGCGGAGTATGGTTTCCGCAAGAAGAGGGCGGCGCTG	398
OY	241	GAAACGTTGCAAGGCTCACAGAAACATCGTGACTTTGATGGAAGTTTACATCACTTT	300
Db	399	GAAACGTTGCAAGGCTCACAGAAACATCGTGACTTTATGGAAGTTTACATCACTTC	458
OY	301	TCTCCAAATGAGCATCACGCTGTCTGTGTTGATGAACCTCGGATGTCAGTGTTCGGAA	360
Db	459	TCTCCCAATGTCATCACGCTGTCTGTGTTGAATCTCTGAGATGTCAGTGTTCGGAA	518
OY	361	TTGCTCTTATATTCGATCACAGGAGTTGTCATGTGATGATACAGCACTTGGCCCGA	420
Db	519	TTGCTTTATATTCGATCATCAAGGAGCTGCTCAATGTGATGATACAGCACTGTGCCGA	578
OY	421	GATGTTTTGGAGGCGCTTGTCTTTCTTCATATAGGGCTATGTCCATGGGGACCTCAAA	480
Db	579	GATGTCTGGAGGCGCTTGTCTTTCTTCATATAGGGCTATGTCCATGAGACTCAAA	638
OY	481	CCAGGTAAACATATTTGTGGAGTGCAGAGATGAATGTTTAACTCATTTGACTTTGACTT	540
Db	639	CCAGGAAACATCTGTGGAGTGCAGAGATGAGTGTATTAAGCTTATTTGACTTTGACTTC	698
OY	541	AGCTTCAAGAAAGGCATCAGAGATGTAAAGTATTTCAAGACAGCGGTATCGGGCTCCA	600
Db	699	AGCTTCAAGAAAGGCATCAGAGATGTAAAGTATTTCAAGACAGCGGTATCGGGCTCT	758
OY	601	GAAACAGATTTGCAAAATTTGCTTGGGCCGAGCTGGCTGCAGATGTATACAGATGTATCC	660
Db	759	GAAACAGAGCTGCAGAACTGCTTGTGGCCAGGCCGAGCTGCAGATGTATACAGATGTATCC	818
OY	661	TCAGCTGTGATCTGTGGAGCTTGAAGATCATTTTACTGGAATGTTCTCGAATGAAA	720
Db	819	TCAGCTGTGATCTGTGGAGCTCGGAATCACTTTACTGGAATGTTCTCGAATGAAAG	878
OY	721	CTGAAACATACAGTCAGATCTCAGAGATGGAAGGCAACACTTCGCTATTAATGATATAC	780
Db	879	CTGAAACATACAGTCAGATCTCAGAGATGGAAGGCAACAGTTCTGCTATTAATGATAT	938
OY	781	ATATTTGCGCAATAAAGCAGTGGTGAATGCGCAATTCGAGCTATCACTTAAGAGACTT	840
Db	939	ATATTTGCGCAATAAAGCAGTGGTGAATGCGCAATTCGAGCTATCACTTAAGAGACTT	998
OY	841	ATCAAAAAGCATGCTTCAATGATATCCAGCAGAAAGAAATTCCTGCTGAATGGAATGTGCG	900
Db	999	ATCAAAAAGCATGCTTCAATGATATCCGAGCAGAAAGAAATTCCTGCTGAAGTGGCATTTGGC	1058
OY	901	AGCCATTTCTTAGCATTTCTTTTGGCCCTCATATTGAAGATCTGTCATCTCCCACT	960
Db	1059	AGCCATTTCTTTAGCATTTCTTTTGGCCCTCATATTGAAGATCTGTCATCTCCCACT	1118
OY	961	CCAAGTCTAAGACTGCTGAATGTGCTGGATGATATTAATCTTGGGAATGAAGAGAAATAT	1020
Db	1119	CCAAGTCTCAAGGCTCTCAATGTGCTGGATGATATTAATCTTGAAGAAATGAAGAGAAATAT	1178
OY	1021	GAAAGATGTTGTAAGATGTATTAAGAGAGAGTGTCAAAAATATGGAACCGAGTGTATCTCTA	1080
Db	1179	GAAAGATGTTGTAAGATGTATTAAGAGAGAGTGTCAAAAATATGGAACCGAGTGTATCTCTG	1238
OY	1081	CTTGTCTCAAGAAATCTTGAGCAGAGCAATGCTTTTGTGTGAGTATGCAAAATGCTGCT	1140

Db	1239	CTGTGTCAAAAGCGAATAATCCTGGCAGAGACAAAGCTTCCGTGTAAGCCAAAGCTTGCT	1298
Oy	1141	GATTCCAAAGCTGCAGCAAAATTA CTGACTGGAAGATGTTTGATGGGAAGTTTGTTGG	1200
Db	1299	GATTCCAAAGCTGCTCAAGAAGTGTGCTGACTGGAGAGATGTTTGAAGGGAAAGTTTGTTGG	1358
Oy	1201	GCTACATTCTACCCGCTGAGTCCCTACAAGAGGGATATCTGTATCAAACTTGCTTTAA	1266
Db	1359	GCTACATTCTAACCGCTGAGTGCTTACAAGAGGGATATCTTTATCAAACTTGCTTTAA	1418
RESULT 2			
LOCUS	AKO30152		
DEFINITION	AKO30152	2846 bp mRNA linear HTC 03-APR-2004	
ACCESSION	AKO30152	Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:493401D07 product:kinase interacting with leukemia-associated gene (etachmin), full insert sequence.	
VERSION	AKO30152.1	GI:26326136	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE			
AUTHORS	Carninci,P. and Hayashizaki,Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
PUBMED	99279253		
REFERENCE	10349636		
AUTHORS			
TITLE	2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
PUBMED	20495374		
REFERENCE	11042159		
AUTHORS			
TITLE	3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Kono,H., Akiyama,J., Nishi,K., Kitsuma,T., Tashiro,H., Itoh,M., Sumi.N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,K., Matsunoto,H., Sakaguchi,S., Ikegami,T., Kasaiwagi,K., Fujikake,K., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawat,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kiru,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
PUBMED	20530913		
REFERENCE	11076861		
AUTHORS			
TITLE	4 THE RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium. Functional annotation of a full-length mouse cDNA collection		
JOURNAL	Nature 409, 685-690 (2001)		
PUBMED	5		
REFERENCE			
AUTHORS	The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.		
TITLE	Analyses of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs		
JOURNAL	Nature 420, 563-573 (2002)		
PUBMED	6 (bases 1 to 2846)		
REFERENCE	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hirooka,T., Hiroyane,T., Horii,F., Imocani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawasaki,J., Kojima,Y., Kondo,S., Kono,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi.K., Nomura,K., Numazaki,R., Ono,M., Ohsato,N., Okazaki,Y., Saito,K., Saich,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shingawa,A., Shiaki,T.,		

FEATURES	COMMENT
<p> TITLE JOURNAL Direct Submission Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Saito-cho, Tsukuba-shi, Ibaraki, Kanagawa 230-0045, Japan (E-mail: genome-gsc.riken.jp, URL: http://genome-gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216) cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome-gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/. Location/Qualifiers 1. .2846 organism="Mus musculus" mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM:DB:4933401D07" /db_xref="taxon:10090" /clone="4933401D07" /sex="male" /cisue_type="testis" /clone_id="RIKEN_full-length enriched mouse cDNA library" /dev_stage="adult" 238..1230 /note="unnamed protein product; kinase interacting with leukemia-associated gene (statman) (MGI:1341908) putative" /codon_start=1 /protein_id="BAC26812.1" /db_xref="GI:26326137" /translation="MTLYGVFTTFSPNPSRCLLELDVSVSELLLYSHOGCSNMIMICHKARDVLEALFLHIEGVADHDKPRNLIANSNECPKLDIFGSPKRGQDVKKIORDGRAPAEONCLAQGLSDPTECAVDLWMSGILILMFSPKMKHIVRSOEBWRNKSAILIDHIFPAKRVYMAAIPAYILRLISMHDHDPGRIPALMAALCPSPFSLPFAHIEDVNLPTPIRLANLVDDVLENDYLEDYEDVDYEDYKESQKQKPVSLVPEKNGRGVFAVRYANAGSQAQKLLTGRMFGKTVAVTFPLSAVKKQYLQTL" 2823..2828 /note="putative" 2846 /note="putative" /note="putative" </p>	<p> Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y. Direct Submission Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Saito-cho, Tsukuba-shi, Ibaraki, Kanagawa 230-0045, Japan (E-mail: genome-gsc.riken.jp, URL: http://genome-gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216) cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome-gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/. Location/Qualifiers 1. .2846 organism="Mus musculus" mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM:DB:4933401D07" /db_xref="taxon:10090" /clone="4933401D07" /sex="male" /cisue_type="testis" /clone_id="RIKEN_full-length enriched mouse cDNA library" /dev_stage="adult" 238..1230 /note="unnamed protein product; kinase interacting with leukemia-associated gene (statman) (MGI:1341908) putative" /codon_start=1 /protein_id="BAC26812.1" /db_xref="GI:26326137" /translation="MTLYGVFTTFSPNPSRCLLELDVSVSELLLYSHOGCSNMIMICHKARDVLEALFLHIEGVADHDKPRNLIANSNECPKLDIFGSPKRGQDVKKIORDGRAPAEONCLAQGLSDPTECAVDLWMSGILILMFSPKMKHIVRSOEBWRNKSAILIDHIFPAKRVYMAAIPAYILRLISMHDHDPGRIPALMAALCPSPFSLPFAHIEDVNLPTPIRLANLVDDVLENDYLEDYEDVDYEDYKESQKQKPVSLVPEKNGRGVFAVRYANAGSQAQKLLTGRMFGKTVAVTFPLSAVKKQYLQTL" 2823..2828 /note="putative" 2846 /note="putative" /note="putative" </p>
<p> ORIGIN polya_signal polya_site polya_site /note="putative" 2846 /note="putative" /note="putative" </p>	
<p> Query Match Best Local Similarity 93.1%; Pred. No. 1.8e-248; Matches 928; Conservative 0; Mismatches 69; Indels 0; Gaps 0; </p>	
<p> QY 264 CATGTCGACTTTGTAATGAGATGTTTACATCCACTTTCTCCAAATGTCGATACGCTG 323 DB 234 CTTTCATCACTTTATATGAGATCTTTTACCAATCACTTCTCTCCAAATGTCGATACGCTG 293 QY 324 TCTGTTGCTGGAACCTCGATGATGCACTGTTTGGGAATGCTCTTATATTCAGATCA 383 DB 294 TCTGCTGCTGGAACCTCGATGATGCACTGTTTGGGAATGCTCTTATATTCAGATCA 353 QY 384 GGGTGTTCATGATGATGATACAGCATTCGCGCCGAGATGTTTGGAGGCGCTTGTCTT 443 DB 354 GGGGTGTCATGATGATGATACAGCATTCGCGCCGAGATGTTTGGAGGCGCTTGTCTT 413 QY 444 TCTTCATCATGAGGCTATATGTCATGCGGACCTCAACCAACGATACATATTTGGAATGC 503 DB 414 TCTTCACCAAGAGGCTATATGTCATGCGGACCTCAACCAACGATACATATTTGGAATGC 473 QY 504 AGGAATGATGTTTAACTCACTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGA 563 DB 474 GGAATGATGATGTTTAACTCACTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGA 533 </p>	

```

QY 564 TGTAAAGTATATTCAGACAGACGGGATCGGGCTCCAGAGACAGAAATGGAAATGGCTT 623
DB 534 CGTAAAGTATATTCAGACAGACGGGATTCGGCTCCGAAAGACAGAGCTGAGAACTGCTT 593
QY 624 GGGCCAGAGCTGGCTGAGAGATACAGAAATGACCTGAGCTGGATGCTGAGACCT 663
DB 594 GGGCCAGAGCTGGCTGAGAGATACAGAAATGACCTGAGCTGGATGCTGAGACCT 653
QY 684 AGGAATCATTTTCTGGAAATGTTCTCAGAAATGAAACATGAAACATGACATGCTCA 743
DB 654 CGGAATCATTTTCTGGAAATGTTCTCAGAAATGAAACATGAAACATGACATGCTCA 713
QY 744 GGAATGGAAGGCAAGAGCTTCCTATTTATTTATGATCAATATTTTCCAGTAAAGAGTGT 803
DB 714 GGAATGGAAGGCAAGAGCTTCCTATTTATTTATGATCAATATTTTCCAGTAAAGAGTGT 773
QY 804 GAATGCCGCAATTCAGAGCTTATCATCTAGAGACCTTATCAAAAGCATGCTTATGATGA 863
DB 774 GAATGCCGCAATTCAGAGCTTATCATCTAGAGACCTTATCAAAAGCATGCTTATGATGA 833
QY 864 TCCAGCAGAGAAATTCCTGCTGATAATGGCATTTGCAAGCCATTCCTTTAGCATTCCTT 923
DB 834 CCGGGCAGAGAAATTCCTGCTGATAATGGCATTTGCAAGCCATTCCTTTAGCATTCCTT 893
QY 924 TGGCCCTCATATTTGAAGATCTGCTCATGCTTCCCACTCCAGTCCAGTCCAGTCCAGT 983
DB 894 TGGCCCTCATATTTGAAGATCTGCTCATGCTTCCCACTCCAGTCCAGTCCAGTCCAGT 953
QY 984 GCTGATGATGATTTATCTTGGAAATGAAAGAAATGAAAGATGTTGTAAGATGTA 1043
DB 954 GCTGATGATGATTTATCTTGGAAATGAAAGAAATGAAAGATGTTGTAAGATGTA 1013
QY 1044 AGAGAGATGTCAGAAATATGAGACAGTGGATCTCACTGTTCCAAAGGAAATCTCTG 1103
DB 1014 AGAGAGATGTCAGAAATATGAGACAGTGGATCTCACTGTTCCAAAGGAAATCTCTG 1073
QY 1104 CAGAGGCAAGTCTTGTGTAGTATGCAATGCTGTGATTCCAAGCTGCGCAGAAAT 1163
DB 1074 CAGAGGCAAGTCTTGTGTAGTATGCAATGCTGTGATTCCAAGCTGCGCAGAAAT 1133
QY 1164 ACTGATCGAAGATGTTGATGAGAAAGTGTGTTGGCTGACATTTACCCGCTGAGTGC 1223
DB 1134 GCTGATCGAAGATGTTGATGAGAAAGTGTGTTGGCTGACATTTACCCGCTGAGTGC 1193
QY 1224 CTACAGAGGAGATCTGTATCAAACTTGTCTTAA 1260
DB 1194 CTACAGAGGAGATCTGTATCAAACTTGTCTTAA 1230

```

```

RESULT 3
LOCUS BM468107 1079 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6432293 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5535653
ACCESSION BM468107
VERSION BM468107.1 GI:18517149
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1079)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be

```

found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLNL123222 row: b column: 06
 High quality sequence stop: 754.
 Location/Qualifiers

FEATURES

source

```

1..1079
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5535653"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_11b="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-Sport6; Site 1: Not1;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."

```

ORIGIN

Query Match 68.4%; Score 861.8; DB 4; Length 1079;
 Best Local Similarity 98.8%; Pred. No. 2,6e-24;
 Matches 888; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

```

QY 267 CGTGAATTTGATGAGAGTGTATACATCCACTTTCTCAATATGTCATACAGCTGTCT 326
DB 1 CGTGAATTTGATGAGAGTGTATACATCCACTTTCTCAATATGTCATACAGCTGTCT 60
QY 327 GTTGCTTGAATCTCTGAGATGTCAGTGTTCGGAATTCCTTATATTCAGTCAAGAGG 386
DB 61 GTTGCTTGAATCTCTGAGATGTCAGTGTTCGGAATTCCTTATATTCAGTCAAGAGG 120
QY 387 TTGTTCCATGTGATGATACAGCATTTGCGCCGAGATGTTTGGAGCCCTTGTCTTCT 446
DB 121 TTGTTCCATGTGATGATACAGCATTTGCGCCGAGATGTTTGGAGCCCTTGTCTTCT 180
QY 447 TCATCATGAGGCTATGTCATGTCAGGACCTCAACACAGTAAACATATGTGGAGTGA 506
DB 181 TCATCATGAGGCTATGTCATGTCAGGACCTCAACACAGTAAACATATGTGGAGTGA 240
QY 507 GAATGAATGTTTAACTCATTTGATCTTGAAGTTCCTTCAAGAGGCAATCAGATGT 566
DB 241 GAATGAATGTTTAACTCATTTGATCTTGAAGTTCCTTCAAGAGGCAATCAGATGT 300
QY 567 AAATGATATTCAGACAGACGGGATTCGAGCTCCAGAGCAGAAATTCGAAATTCGTTG 626
DB 301 AAATGATATTCAGACAGACGGGATTCGAGCTCCAGAGCAGAAATTCGAAATTCGTTG 360
QY 627 CAGAGCTGGCTGAGAGGATGATGAAATGTAACCTCAGCTGTGATCTGTGGAGCTTGA 686
DB 361 CAGAGCTGGCTGAGAGGATGATGAAATGTAACCTCAGCTGTGATCTGTGGAGCTTGA 420
QY 687 AATCATTTTACTGGAATGTTCTCAGAAATGAACTGAAACATACAGTCAGATCTCAGGA 746
DB 421 AATCATTTTACTGGAATGTTCTCAGAAATGAACTGAAACATACAGTCAGATCTCAGGA 480
QY 747 ATGGAAGGCAAAAGCTTCTGCTATTTATGATCAATTTTGGCAGTAAAGAGTGTGAA 806
DB 481 ATGGAAGGCAAAAGCTTCTGCTATTTATGATCAATTTTGGCAGTAAAGAGTGTGAA 540
QY 807 TGGCGCAATTCAGGCTATCACCTTAAGAGACCTTATCAAAAGCAAGCTTCATGATGATCC 866
DB 541 TGGCGCAATTCAGGCTATCACCTTAAGAGACCTTATCAAAAGCAAGCTTCATGATGATCC 600
QY 867 AAGCAGAAAGAAATTCCTGCTGAAATGGAATGTCAGAGCCATTCCTTACATTCCTTTGC 926
DB 601 AAGCAGAAAGAAATTCCTGCTGAAATGGAATGTCAGAGCCATTCCTTACATTCCTTTGC 660
QY 927 CCTCATATTTGAAGATCTGCTGATCTTCCCATCTCAGTCTTACAGTCTGATGTGCT 986
DB 661 CCTCATATTTGAAGATCTGCTGATCTTCCCATCTCAGTCTTACAGTCTGATGTGCT 720
QY 987 GGATGATGATTTATCTTGGAAATGGAAGAAATATGAATGATGTAAGAAATGTAAGA 1046
DB 721 GGATGATGATTTATCTTGGAAATGGAAGAAATATGAATGATGTAAGAAATGTAAGA 780

```

QY 1047 GGAGTGTCAAAATATGACACGATGATCTTACTTGTTCAGAAAGAAAATCTGACG 1106
DB 781 GGAGTGTCAAAATATGACACGATGATCTTACTTGTTCAGAAAGAAAATCTGACG 840
QY 1107 AGGACAAAGTC-TTGTGTGAGTATGACAAATGCTG-TGTATTCGAAAGCTGGCGAATAAT 1163
DB 841 AGGACAAAGTC-TTGTGTGAGTATGACAAATGCTG-TGTATTCGAAAGCTGGCGAATAAT 899

RESULT 4
BM451184 917 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCY 6392719 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5495159
DEFINITION 5', mRNA sequence.
ACCESSION BM451184
VERSION BM451184.1 GI:18500224
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 917)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1M12121 row: j column: 24
High quality sequence stop: 718.
Location/Qualifiers
1..917
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5495159"
/issue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_67"
/note="Organ: eye; Vector: pCMV-SPORT6; Site: 1: NciI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."

ORIGIN

Query Match 57.4%; Score 723.8; DB 4; Length 917;
Best Local Similarity 99.7%; Pred. No. 8.7e-201;
Matches 725; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 534 TGGACTTACCTTAAAGAGGCAATCAGATGTAAGTATATTCACAGACGCGATTCG 593
DB 21 TGGGCTTAGCTTAAAGAGGCAATCAGATGTAAGTATATTCACAGACGCGATTCG 80
QY 594 GGCTCCAGAGCAGAAATGCAAAATGCTGGCCAGGCTGGCCCTGACAGATGATTCAG 653
DB 81 GGCTCCAGAGCAGAAATGCAAAATGCTGGCCAGGCTGGCCCTGACAGATGATTCAG 140
QY 654 ATGTACCTCAGCTGTGATCTGTGAGCCCTAGGAATCATTTTACGAAATGTTCTCAG 713
DB 141 ATGTACCTCAGCTGTGATCTGTGAGCCCTAGGAATCATTTTACGAAATGTTCTCAG 200
QY 714 AATGAAATGAAACATTCAGTCTGAGATTCAGAAATGAAAGCAACAGTTCTCTATTA 773
DB 201 AATGAAATGAAACATTCAGTCTGAGATTCAGAAATGAAAGCAACAGTTCTCTATTA 260
QY 774 TGATCACAATTTTGCAGTAAGAGAGTGTGATGCGCAATTCAGCCTATCACTTAAG 833

DB 261 TGATCACAATTTTGCAGTAAGAGAGTGTGATGCGCAATTCAGCCTATCACTTAAG 320
QY 834 AGACCTTATCAAAAGCAGCTTCATGATGATTCAGACAGAAATTCCTGCGAATGCG 893
DB 321 AGACCTTATCAAAAGCAGCTTCATGATGATTCAGACAGAAATTCCTGCGAATGCG 380
QY 894 ATTGTGACAGCCATTCTTATGATTCCTTTTGCCTCATATATGAAAGATCTGTCATGCT 953
DB 381 ATTGTGACAGCCATTCTTATGATTCCTTTTGCCTCATATATGAAAGATCTGTCATGCT 440
QY 954 TCCCACTCCAGTCTTAAAGCTGCTGAATGCTGATGATGATTTATCTTGGAAATGAGA 1013
DB 441 TCCCACTCCAGTCTTAAAGCTGCTGAATGCTGATGATGATTTATCTTGGAAATGAGA 500
QY 1014 GGAATATGAAAGATGTTGTGAGATGCTTAAAGAGAGCTCAAAATATGACCCAGTGT 1073
DB 501 GGAATATGAAAGATGTTGTGAGATGCTTAAAGAGAGCTCAAAATATGACCCAGTGT 560
QY 1074 ATCTCTACTTGTTCAGAAAGAAATCTGCGAGAGACAAAGCTTGTGTGATGATGACAA 1133
DB 561 ATCTCTACTTGTTCAGAAAGAAATCTGCGAGAGACAAAGCTTGTGTGATGATGACAA 620
QY 1134 TCTGTGATGATTCAGAAAGCTGCGAGAAATCTGATGATGATGATGATGATGATGAT 1193
DB 621 TCTGTGATGATTCAGAAAGCTGCGAGAAATCTGATGATGATGATGATGATGATGAT 680
QY 1194 TGTGTGTGATGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1253
DB 681 TGTGTGTGATGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 740
QY 1254 GCTTTAA 1260
DB 741 GCTTTAA 747

RESULT 5
BY712344
LOCUS BY712344
DEFINITION RIKEN full-length enriched, 10, 11 days embryo whole body
MUS musculus cDNA clone 281045404 5', mRNA sequence.
ACCESSION BY712344
VERSION BY712344.1 GI:27123621
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 1015)
AUTHORS Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Oshio, N., Saito, R., Suzuki, H., Yamana, A.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schumacher, C., Gotohori, T., Baldarelli, R., Hill, D.P., Bulc, C.,
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,
Chochia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
Fletcher, C.F., Forrest, A., Fraser, K.S., Gaasterland, T.,
Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, B.D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
Kurachkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Meglert, D.R.,
Maltre, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okido, T., Pavani, W.J., Petrea, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J.V., Q.D., Ramchandran, S.,
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shinada, K.,
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
Wells, C., Wilming, L.G., Wyshew-Boris, A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,

Itch, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
Shinagawa, A., Yasuniishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

JOURNAL
MEDLINE
PUBMED
12466851
22354683

CONTACT: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.jp, url: http://genome.gsc.riken.jp/

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P.,
Fukuda, S., Hoshizume, W., Hayashida, K., Hirozane, T., Hori, F.,
Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y.,
Kondo, S., Kono, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M.,
Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N.,
Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,
Takeda, Y., Waki, K., Watanabe, A., Watanabe, M. and Hayashizaki, Y.
Direct Substitution
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedias: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.

FEATURES

source
Location/Qualifiers
1..1015
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="2810454004"
/issue_type="whole body"
/dev_stage="10, 11 days embryo"
/clone_lib="RIKEN full-length enriched, 10, 11 days embryo
whole body"

ORIGIN

Query Match 49.6%; Score 624.4; DB 6; Length 1015;
Best Local Similarity 90.2%; Pred. No. 1.4e-171;
Matches 679; Conservative 0; Mismatches 72; Indels 2; Gaps 1;
1 ATGGCGGAGATCCGGCTGCGGCGGCGGAGAGCCCGCTTTCTGGAGGCTTCGGG 60
159 ATGGCGGAGATCCGGCTGCGGCGGCGGAGAGCCCGCTTTCTGGAGGCTTCGGG 218
61 CGGCTGTGACAGGTACAGAGCCGCTGTGGTAGGCGGCTTCGCTCGGTATCGGGTT 120
219 CGGCTGTGACAGGTACAGAGCCGCTGTGGTAGGCGGCTTCGCTCGGTATCGGGTT 278
121 CGCTGTGCGGCAACCTGTGCTGCGGCGGCGGCTTCAGAGAGTTCTTGGCCGAGGA 180
279 CGGTGTGCGGTAACCTGTGCTGCGGCGGCGGCTTCAGAGAGTTCTTGGCCGAGGA 338
181 ACCACCGGCGGCTGCGGCTTCGCTGCGGCGGAGTATGTTTCCGCAAGAGGCGGCGCTG 240
|||||

Db 339 ACCACCGGCGGCTGCGGCTTCGCGGCGGAGAGTATGTTTCCGCAAGAGGCGGCGCTG 398
Qy 241 GAACAGTTGACAGGTACAGAGCAATCGTGAATTTGATGAGTGTATTAATCACTTT 300
Db 399 GANCAAGTTGACAGGTACAGAGCAATCGTGAATTTGATGAGTGTATTAATCACTTT 458
Qy 301 TCTCCAAATGTCATCAACGCTGTCTGTGCTTGAATCTTGGATGTCAGTGTTCGGAA 360
Db 459 TCTCCAAATGTCATCAACGCTGTCTGTGCTTGAATCTTGGATGTCAGTGTTCGGAA 518
Qy 361 TTGCTTTATATTCAGTACACAGAGTTGTTCCAGTGAATGATACAGATTTGGCCGGA 420
Db 519 TTGCTTTATATTCAGTACATAGAGGCTGCTTCAGTGAATGATACAGATTTGGCCGGA 578
Qy 421 GATGTTTGAAGGCTGCTTCTTCTCATCATGAGGCTATGTCATGCGGACCTCAA 480
Db 579 GATGTTTGAAGGCTGCTTCTTCTCATCATGAGGCTATGTCATGCGGACCTCAA 638
Qy 481 CCACGTACATATTTGAGTGCAGAGTAATGATTTTAACTCATTTGACTTTGACTT 540
Db 639 CCACGTACATATTTGAGTGCAGAGTAATGATTTTAACTCATTTGACTTTGACTT 698
Qy 541 ACCTTCAAGAGGCAATCAGATGTTAAAGTATTTACAGACAGAGGATGCGGCTCA 600
Db 699 ACCTTCAAGAGGCAATCAGATGTTAAAGTATTTACAGACAGAGGATGCGGCTCT 758
Qy 601 GAAGCAATTTGCAAAATTTGCTGGCCGAGCTGCTGCAGAGTATCAAGATGTAAC 660
Db 759 GAAGCAATTTGCAAAATTTGCTGGCCGAGCTGCTGCAGAGTATCAAGATGTAAC 818
Qy 661 TGAAGTTGATCTGTGAGCTTACAGATTTTATCTGGAATGTTCTCAGGAATGAA 720
Db 819 TGAGCTTTGATCTGTGAGCTTACAGATTTTATCTGGAATGTTCTCAGGAATGAA 878
Qy 721 CTGAACATACAGTCAAGATTTCAAGAAATGGAAG 753
Db 879 CTGAACATACAGTCAAGATTTCAAGAAATGGAAG 909
|||||

RESULT 6
BG547675 657 bp mRNA linear EST 04-APR-2001
LOCUS 602575520P1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4703645 5,
DEFINITION
mRNA sequence.
ACCESSION BG547675
VERSION BG547675.1 GI:13546340
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: CLOUTEC Laboratories, Inc.
cDNA Library Preparation: CLOUTEC Laboratories, Inc.
DNA Sequencing by: InCyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/BLN at:
http://image.llnl.gov
Plate: L10M1542 row: 9 column: 06
High quality sequence stop: 657.

FEATURES

source
Location/Qualifiers
1..657
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4703645"
/lab_host="DH10B (TI phage-resistant)"

/clone.lib="NIH_MGC_77"
/note="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1:
SfiI (ggcgctcgcc); Site 2: SfiI (ggcattatggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGCGCATTAATGCGC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGCGCGCGCATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.9
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 49.1%; Score 619.2; DB 4; Length 657;
Best Local Similarity 99.1%; Pred. No. 4.2e-170; Mismatches 3; Indels 3; Gaps 3;
Matches 654; Conservative 0; Mismatches 3; Indels 3; Gaps 3;

QY 577 CAGACAGACGGGTATCGGGCTCCAGAGCAGAAATTCGCTTGGCCAGGCTGCG 636
DB 1 CAGACAGACGGGTATCGGGCTCCAGAGCAGAAATTCGCTTGGCCAGGCTGCG 60
QY 637 CTGCAAGATGATACAGATGATCTTACCTGCTGCTGATCTGAGCCTGAGATCATTTTA 696
DB 61 CTGCAAGATGATACAGATGATCTTACCTGCTGCTGATCTGAGCCTGAGATCATTTTA 120
QY 697 CTGGAATGTTCTCAGGAATGAATGAACATACAGTGCATCTCAGGAATGGAAGGCA 756
DB 121 CTGGAATGTTCTCAGGAATGAATGAACATACAGTGCATCTCAGGAATGGAAGGCA 180
QY 757 AACAGTTCGCTATTATTGATCATATTTCAGAGTAAGCAGTGTGATTCGCGCAAT 816
DB 181 AACAGTTCGCTATTATTGATCATATTTCAGAGTAAGCAGTGTGATTCGCGCAAT 240
QY 817 CCAGCTATCACTAAGAGCTTATCAAAAGCATGCTTATGATGATCCAGCAGAGA 876
DB 241 CCAGCTATCACTAAGAGCTTATCAAAAGCATGCTTATGATGATCCAGCAGAGA 300
QY 877 ATTCCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 936
DB 301 ATTCCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
QY 937 GAAGATCTGATCTTCCACTCCAGTGTCTAAGACTGCTGATGATGATGATGAT 996
DB 361 GAAGATCTGATCTTCCACTCCAGTGTCTAAGACTGCTGATGATGATGATGAT 420
QY 997 TATCTTGGAGTGAAGAGAAATGAAGATGTTGTAGAGATGTAAGAGAGTGTCAA 1056
DB 421 TATCTTGGAGTGAAGAGAAATGAAGATGTTGTAGAGATGTAAGAGAGTGTCAA 480
QY 1057 AAATAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1116
DB 481 AAATAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 1117 TTGTGATGATCAAAATGCTGTGATTCGAAAGCTGCGCAGAAATTAATGATGAGG 1176
DB 541 TTGTGATGATCAAAATGCTGTGATTCGAAAGCTGCGCAGAAATTAATGATGAGG 599
QY 1177 ATGTTGATGGAAGTTTGTGTGCTACATTTCACTGAGTGCCTTCAAGAGGGGA 1236
DB 600 ATGTTGATGGAAGTTTGTGTGCTACATTTCACTGAGTGCCTTCAAGAGGGGA 657

RESULT 7
BE895119 899 bp mRNA linear EST 20-OCT-2000
LOCUS 601436041F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3920939 5',
DEFINITION mRNA sequence.
ACCESSION BE895119
VERSION BE895119.1 GI:10358193
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
1 (bases 1 to 899)
NIH-MGC <http://mgs.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: ATCC/DCPD/BTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
plate: LLN9753 row: b column: 12
High quality sequence stop: 626.
Location/Qualifiers

FEATURES
source
1..899
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3920939"
/issue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone.lib="NIH_MGC_72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."

ORIGIN

Query Match 49.0%; Score 617.8; DB 2; Length 899;
Best Local Similarity 99.7%; Pred. No. 1.2e-169; Mismatches 2; Indels 0; Gaps 0;
Matches 619; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 337 CTCCTGATGTCAGTGTTCGGAATTCCTTATATTCAGTCACCGAGTTGTCATG 396
DB 1 CTCCTGATGTCAGTGTTCGGAATTCCTTATATTCAGTCACCGAGTTGTCATG 60
QY 397 TGGATATACAGCATTTGCGCCGAGATGTTTGGAGGCCCTTCTTCTCATCATGAG 456
DB 61 TGGATATACAGCATTTGCGCCGAGATGTTTGGAGGCCCTTCTTCTCATCATGAG 120
QY 457 GGCTATGTCAGTGGAGCTCAACACGTAACATTTGTGAGTGCAGAGATGAATGT 516
DB 121 GGCTATGTCAGTGGAGCTCAACACGTAACATTTGTGAGTGCAGAGATGAATGT 180
QY 517 TTTAACTGATGACTTTGGAATTGCTTCAAGAGCAATCAGGATGTAAGTATAT 576
DB 181 TTTAACTGATGACTTTGGAATTGCTTCAAGAGCAATCAGGATGTAAGTATAT 240
QY 577 CAGACAGACGGGTATCGGGCTCCAGAGCAGAAATTCGCTTGGCCAGGCTGCG 636
DB 241 CAGACAGACGGGTATCGGGCTCCAGAGCAGAAATTCGCTTGGCCAGGCTGCG 300
QY 637 CTGCAAGATGATACAGATGATCTTACCTGCTGCTGATCTGAGCCTGAGATCATTTTA 696
DB 301 CTGCAAGATGATACAGATGATCTTACCTGCTGCTGATCTGAGCCTGAGATCATTTTA 360
QY 697 CTGGAATGTTCTCAGGAATGAATGAACATACAGTGCATCTCAGGAATGGAAGGCA 756
DB 361 CTGGAATGTTCTCAGGAATGAATGAACATACAGTGCATCTCAGGAATGGAAGGCA 420
QY 757 AACAGTTCGCTATTATTGATCATATTTCGAGTAAGCAGTGTGATTCGCGCAAT 816
DB 421 AACAGTTCGCTATTATTGATCATATTTCGAGTAAGCAGTGTGATTCGCGCAAT 480
QY 817 CCAGCTATCACTAAGAGCTTATCAAAAGCATGCTCATGATGATCCAGCAGAGA 876
DB 481 CCAGCTATCACTAAGAGCTTATCAAAAGCATGCTCATGATGATCCAGCAGAGA 540
QY 877 ATTCCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 936

```

Db      541 ATTCTGCTGAATGAGCATTTGTCAGCCCATCTTATGACATCTTTTGCCCTCATATT 600
QY      937 |||||||
Db      601 GAGATCTGCTCAGCTTTC 621

RESULT 8
B1562420      772 bp      mRNA      linear      EST 05-SEP-2001
B1562420      60326316P1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5298674 5',
LOCUS      mRNA sequence.
DEFINITION      B1562420.1 GI:15449747
ACCESSION      B1562420
VERSION      B1562420
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE      1 (bases 1 to 772)
AUTHORS      NIH-MGC http://img.ncbi.nlm.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgabbs-r@mail.nih.gov
              Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
              cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
              Toshikuni and Piero Carninci (RIKEN)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLML1755 row: P column: 03
              High quality sequence scop: 727.
              Location/Qualifiers
                1..772
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:5298674"
                /lab_host="DH10B"
                /clone_lib="NIH_MGC_97"
                /note="Organ: testis; Vector: pBluescript (modified
                pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
                (gtcagc); Oligo-dT primed using primer
                5'-TTTTTTTTTTTTTNN-3', size-selected for average
                insert size 2.2 kb and normalized to R01 5. This is a
                primary library enriched for full-length clones and
                constructed using the Cap-trapper method (Carninci, in
                preparation). Library constructed by M. Brownstein
                (NHGRI/NHGRI, National Institutes of Health). Note: this is
                a NIH_MGC library."

ORIGIN
Query Match      48.9%; Score 616.4; DB 4; Length 772;
Best Local Similarity 99.4%; Pred. No. 2.9e-169;
Matches 628; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY      274 TTGTATGAGAGTGTTCACATCACTTTCTCCAAATGTCGATCAGCTGTCTGTTGCTT 333
Db      108 TTGTATGAGAGTGTTCACATCACTTTCTCCAAATGTCGATCAGCTGTCTGTTGCTT 167
QY      334 GAACCTCTGGATGTCAGTGTGGGAATGCTCTATATTCGATCAGCAGCGGTTGCTC 393
Db      168 GAACCTCTGGATGTCAGTGTGGGAATGCTCTATATTCGATCAGCAGCGGTTGCTC 227
QY      394 ATGTGATGATACAGATTGCGCCCGAGATGTTTGGAGGCGCTTCTTTCATCAT 453
Db      228 ATGTGATGATACAGATTGCGCCCGAGATGTTTGGAGGCGCTTCTTTCATCAT 287
QY      454 GAGGCTATGTCCATGCGGAGCTCAACCGTACATATTTGGAGTGCAGAGATGAA 513

```

```

Db      288 GAGGCTATGTCCATGCGGAGCTCAACACGTAACATATTGTGAGTGCAGAGATGAA 347
QY      514 TGTTTAACTCATTTGACTTTTGGACTTTAGCTTCAAGAGAGCATAGGATGTAAGAT 573
Db      348 TGTTTAACTCATTTGACTTTTGGACTTTAGCTTCAAGAGAGCATAGGATGTAAGAT 407
QY      574 ATTCAGACAGACGGGATGCGGCTCCAGAGAGAAATTCGAAATTTCTTGCCCGAGCT 633
Db      408 ATTCAGACAGACGGGATGCGGCTCCAGAGAGAAATTCGAAATTTCTTGCCCGAGCT 467
QY      634 GGGCTGCAGAGATATCAGAAATGATACCTGAGCTGTGATCTGTGAGCCTTAGAATCAT 693
Db      468 GGGCTGCAGAGATATCAGAAATGATACCTGAGCTGTGATCTGTGAGCCTTAGAATCAT 527
QY      694 TTACTGAAATGTTCTCAGAAATGAATGAACATACATACAGATCTCAGAAATGAAG 753
Db      528 TTTCTGAAATGTTCTCAGAAATGAATGAACATACATACAGATCTCAGAAATGAAG 587
QY      754 GCACACAGTTCTGCTATTTATGATACA-TATTTGCGAGTAAAGCAGTGTGAATGCC 812
Db      588 GCACACAGTTCTGCTATTTATGATACA-TATTTGCGAGTAAAGCAGTGTGAATGCC 647
QY      813 AATTCAGGCTATCAGCTTAAGAGACCTTATCAAAAGCATGCTTATGATCCAGACG 872
Db      648 AATTCAGGCTATCAGCTTAAGAGACCTTATCAAAAGCATGCTTATGATCCAGACG 707
QY      873 AAGAAATTCCTGCTGAATGCAATGTCAGCC 904
Db      708 AAGAAATTCCTGCTGAATGCAATGTCAGCC 739

RESULT 9
B0541306      932 bp      mRNA      linear      EST 13-SEP-2002
B0541306      AGENCOURT_10327419 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6572970
LOCUS      5', mRNA sequence.
DEFINITION      B0541306
ACCESSION      B0541306
VERSION      B0541306.1 GI:22851747
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE      1 (bases 1 to 932)
AUTHORS      NIH-MGC http://img.ncbi.nlm.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgabbs-r@mail.nih.gov
              Tissue Procurement: DCTD/DRP
              cDNA Library Preparation: Rubin Laboratory
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: L1CM2765 row: O column: 18
              High quality sequence scop: 631.
              Location/Qualifiers
                1..932
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:6572970"
                /rfeature_type="carcinoma, cell line"
                /lab_host="DH10B (phage-resistant)"
                /clone_lib="NIH_MGC_40"
                /note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
                Site 2: EcoRI; cDNA made by oligo-dT priming.
                Directionally cloned into EcoRI/XhoI sites using the
                following 5' adaptor: GGCACGAG(G). Library constructed by
                Ling Hong in the laboratory of Gerald M. Rubin (University
                of California, Berkeley) using Zap-cDNA synthesis kit

```


RESULT	10
B1259568	
LOCUS	
DEFINITION	B1259568 875 bp mRNA linear EST 17-JUL-2001 60299683301 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5108094 5', mRNA sequence.
ACCESSION	B1259568
VERSION	B1259568.1 GI:14817032
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OY	297	CTTTCTCCAAATGTCATCAAGCTGCTGTGCTTGAATCCCTGATGTGACAGTTC	356
Db	1	CTTTTCTCCAAATGTCATCAAGCTGCTGTGCTTGAATCCCTGATGTGAC - GTTTC	59
OY	357	GGAATGTGCTTATATATTCAGTCAACAAGGGTGTTCATGTGATGATACAGATTGGC	416
Db	60	GGAATGTGCTTATATATTCAGTCAACA - GGTGTTCATGTGATGATACAGCA - TGTGC	117
OY	417	CCGAGATGTTTGGAGGCCCTTGTCTTTCTTCATCATGAGGGCTATGTCCATGGGAACT	476
Db	118	CCGAGATGTTT - GGAGGCCCTGTCTTTCTTCATCATGAGGGCTATGTCCATGGGAACT	175
OY	477	CAAAACAAGTACATATTTGAGAGTCGACAGAAATGAATGTTTAACTCATATGACTTTGG	536
Db	176	CAAAACAAGTACATAT - TGTGAGTGCAGAAATGAATGTTTAACTCATATGAC - TTGG	233
OY	537	ACTTAGCTTCAAAGAGGCATCAGATGTAAAGTAAATTCAGACAGACGGGATACGGGC	596
Db	234	ACTTAGCTTCAAAAGAGGCATCAGATGTAAAGTAAATTCAGACAGACGGGATACGGGC	293
OY	597	TCCGAAGACGAATTCGAAAATTCGTGGCCCAAGCTGGCCCTGCAGATGATACGAATG	656
Db	294	TCCGAAGACGAATTCGAAAATTCGTGGCCCAAGCTGGCCCTGCAGATGATACGAATG	353
OY	657	TACCTCAGCTGTGATCTGTGAGACCTAGGAATATTTAATGGAAATGTTCTCAGAAAT	716
Db	354	TACCTCAGCTGTGATCTGTGAGACCTAGGAATCATTTTACTGGAAATGTTCTCAGAAAT	413
OY	717	GAACTGAACATACAGTCAGATCTCAGGAATGGAAGCAAAAGTTCTGCTATATTGA	776
Db	414	GAACTGAACATACAGTCAGATCTCAGGAATGGAAGCAAAAGTTCTGCTATATTGA	473
OY	777	TCACATATTTGCCAGTAAAGCAGTGTGATTCGGCAATTCAGGCTATCACTTAAGAGA	836
Db	474	TCACATATTTGCCAGTAAAGCAGTGTGATTCGGCAATTCAGGCTATCACTTAAGAGA	533
OY	837	CCTTATCAAAAGCATGCTTCAATGATGATCCAAAGAGAAATTCGTGTGAAATGGCAAT	896
Db	534	CCTTATCAAAAGCATGCTTCAATGATGATCCAAAGAGAAATTCGTGTGAAATGGCA - T	592

Qy	897	GTGACGCCCAATCTTTAGCATTCCTTTGGCCCTCAATTTGAAATCT-GGTCATGCTTC	955
Db	593	GTGACGCCCAATCTTTAGCATTCCTTTGGCCCTCAATTTGAAATCTGGGTCAATGCTTC	652
Qy	956	CCAATCCAGTC-CTAAGACTGCTGAATGTGCTGGATGATGATTAATCTTGGGAATGAAG	1012
Db	653	CCAATCCAGGGCTAAAGACTGCTGAAGTGTCTGGATGATGATTAATCTTGGGAATGAAGC	712
Qy	1013	AGGAATATGAAG-ATGTGTAGAAG--ATGTAAAGAGAGTGTCAAAATATATGACACCAG	1065
Db	713	AGGAATATGAAGATGTTGTAGACGGAATGTAAAGAGAGATGTCAACCAATTTGGACCA	772
Qy	1070	TGTATCTCTACTGT-----CCAAAGGAAATCCTGGCAGAGGACA-AGTCTTTGTT	1122
Db	773	GGTGATATTCCTCTACTTGTGTCCAAAGGAAAAATCTCGGGCGAGGACACAGTCTTTGTA	832
Qy	1123	G-AGTATGCAAAAT-GCTGTGATTCCAAAGCTGGCGCAGAAAT	1163
Db	833	GAAATATGCAAAATGTTGGGATTCCAAAGATGGCCGCAAAATTT	875

RESULT 11	LOCUS	DEFINITION
CF593363	834 bp	mRNA
CF593363	15622279	NIH_MGC_147 Homo sapiens
CF593363	30531268	5' mRNA sequence.

VERSION	CF593363.1	GI:36346826
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	

REFERENCE	1 (bases 1 to 834)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Daniela S. Gerhard, Ph.D.

Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MCC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.lnl.gov>
Plate: NDAM621 row: m column: 05
High quality sequence stop: 660.

```

FEATURES
    source
        location/Qualifiers
            1. 834
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:30531268"
                /tissue_type="Human Placenta"
                /lab_host="DH10B Tona"
                /clone_1fb="NIH MGC 147"
                /note="Organ: placenta; Vector: pbluescript; Site 1:
                all-XhoI; Site 2: BamH; Oligo-dT primed using primer
                5'-TTTTTTTTTTTTT-3', size-selected for average
                insert size 2.3 kb and normalized to RQF 5. This is a
                primary library enriched for full-length clones and
                constructed using the Cap-trapper method (Carninci, in
                preparation). Library constructed by M. Brownstein
                (NIH/MHGRI, National Institutes of Health). Note: This is
                a NIH MGC library."

```

Query Match	47.1%	Score 593.2;	DB 7;	Length 834;
Best Local Similarity	97.6%;	Pred. No. 2e-162;		
Matches 645;	Conservative 0;	Mismatches 9;	Indels 7;	Gaps 4;

Oy	1	ATGGCGGGAAATCCGGCTAGCGCTGGGGGGCGGAGCGCGCGGTTTCTGAGAGGCTTCGGG	60
Db	155	ATGGCGGGAAATCCGGCTAGCGCTGGGGGGCGGAGCGCGCGGCTTTCTGAGAGGCTTCGGG	214
Oy	61	CGGCTGTGGCAGGTATCAGAGCCGTCGTGGGTATGCGGCTCTCCGCTCGGTGTATCGGGATT	120
Db	215	CGGCTGTGGCAGGTATCAGAGCCGTCGTGGGTATGCGGCTCTCCGCTCGGTGTATCGGGATT	274
Oy	121	CGCTGCTGGCGCAACCTTGCGTCGGCCCCCGGGGCGCTTCAAGCAGTTCTTGGCGCCAGGA	180
Db	275	CGCTGCTGGCGCAACCTTGCGTCGGCCCCCGGGGCGCTTCAAGCAGTTCTTGGCGCCAGGA	334
Oy	181	ACCAACCGGGGCTGGGGGCTCTGCGCGCGAGTATGTTCCGCAAGAGAGGGGCGGCGCTG	240
Db	335	ACCAACCGGGGCTGGGGGCTCTGCGCGCGAGTATGTTCCGCAAGAGAGGGGCGGCGCTG	394
Oy	241	GAACAGTTTGCAGGGTCAACAGAAACATGTGTACTTTGTATGGAGTGTTTTACATCCACTTT	300
Db	395	GAACAGTTTGCAGGGTCAACAGAAACATGTGTACTTTGTATGGAGTGTTTTACATCCACTTT	454
Oy	301	TCTCAATATGAGCATACGCTGTCTGTGGCTTGAACTCCGCGATGTCAAGTGTTCGAAA	360
Db	455	TCTCAATATGAGCATACGCTGTCTGTGGCTTGAACTCCGCGATGTCAAGTGTTCGAAA	514
Oy	361	TTGCTCTTAATATTCACAGTCAACAGGGTGTTCATGTGGATGATACACATTCGCGCCGA	420
Db	515	TTGCTCTTAATATTCACAGTCAACAGGGTGTTCATGTGGATGATACACATTCGCGCCGA	574
Oy	421	GATGTTTTGGAGGCGCTTGCTTTTCTCATCATGAGGGCTATGTCCATGGCGGACTCAAA	480
Db	575	GATGTTTTGGAGGCGCTTGCTTTTCTCATCATGAGGGCTATGTCCATGGCGGACTCAAA	634
Oy	481	CCAAGTAACTATTTGTGGAGTGCAGAGATGAATGTTTAACTCATTTGACTTTGA	540
Db	635	CCAAGTAACTATTTGTGGAGTGCAGAGATGAATGTTTAACTCATTTGACTTTGA	694
Oy	541	A-GCTTCAAGAAAGGCATACAGATGTAAAGTAAATTTACACAGACGCGGTATC-GGGCTC	598
Db	695	AGGCTTCAAGAAAGGCATACAGATGTAAAGTAAATTTACACAGACGCGGTATC-GGGGCTC	754
Oy	599	CAGAAAGCAATTTGCAAAATTTGCTT--GGGCCAGGCT--GGCTGCAAGTGTATACAGA	653
Db	755	CAGAAAGCAATTTGCAAAATTTGCTTGGGCCCGAGGCTTGACAGTGTATACAGA	814
Oy	654	A 654	
Db	815	A 815	

RESULT 12	
B133581	
LOCUS	682 bp mRNA Linear EST 30-Jul-2001
DEFINITION	M02997065F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5139184 5' ,
ACCESSION	B133581
VERSION	B133581
KEYWORDS	B133581.1 GI:15018238
SOURCE	EST.
ORGANISM	Homo sapiens (human)
	Homo sapiens
	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	1 (bases 1 to 682)
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished (1999)
	Contact: Robert Strausberg, Ph.D.
	Email: cgabds-i@mail.nih.gov
	Tissue Procurement: ATCC
	cDNA Library Preparation: Life Technologies, Inc.

CDNA library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLNL1342 row: j column: 17
High quality sequence stop: 674.

FEATURES

source

1..682
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5139184"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_11b="NIH_MGC_12"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

ORIGIN

Query Match 46.4%; Score 585.2; DB 4; Length 682;
Best Local Similarity 97.3%; Pred. No. 4.1e-160;
Matches 648; Conservative 0; Mismatches 13; Indels 5; Gaps 5;

QY 1 ATGGCGGGATCCGGCTGGCGGGCGGGAGACCGCGCTTTCTGGAGGCGCTTCGGG 60
DB 19 ATGGCGGGATCCGGCTGGCGGGCGGGAGACCGCGCTTTCTGGAGGCGCTTCGGG 78
QY 61 CGGCTGTGCGAGGTACAGAGCCGTCTGGGTAGCGGCTCTCCGCTCGGTATTCGGGTT 120
DB 79 CGGCTGTGCGAGGTACAGAGCCGTCTGGGTAGCGGCTCTCCGCTCGGTATTCGGGTT 138
QY 121 CGCTGTGCGAGACCTTGGCTGCGCCCGCGGCGCTTCAAGAGTTCTTCCGCGAGAA 180
DB 139 CGCTGTGCGAGACCTTGGCTGCGCCCGCGGCGCTTCAAGAGTTCTTCCGCGAGAA 198
QY 181 ACCACGGGGGCTGGCGCTCTGGCGGGAGTATGTTCCGCAAGAGAGGGCGGCGTG 240
DB 199 ACCACGGGGGCTGGCGCTCTGGCGGGAGTATGTTCCGCAAGAGAGGGCGGCGTG 258
QY 241 GAACAGTTGCGAGGTACAGAAACATCGTACCTTTGTATGAGTGTTCATCACTTT 300
DB 259 GAACAGTTGCGAGGTACAGAAACATCGTACCTTTGTATGAGTGTTCATCACTTT 318
QY 301 TCTCAATGTGTCATCAGCTGTCTGTGCTTGAATCTCTGATGTCAAGTCTTTCGAA 360
DB 319 TCTCAATGTGTCATCAGCTGTCTGTGCTTGAATCTCTGATGTCAAGTCTTTCGAA 378
QY 361 TTGCTTTTATTCAGTACACGAGGTGTTCATGTGATGATTAAGCATTTGCCCGCA 420
DB 379 TTGCTTTTATTCAGTACACGAGGTGTTCATGTGATGATTAAGCATTTGCCCGCA 438
QY 421 GATGTTTGGAGGCGCTTGTCTTTCTTCATCATGAGGCTATGTCCATGCGAGCTCAAA 480
DB 439 GATG-TTTGGAGGCGCTTGTCTTTCTTCATCATGAGGCTATGTCCATGCGAGCTCAAA 497
QY 481 CCAAGTAAATTTGAGAGTGAAGAAATGTTTAACTCAATGACTTGGACTT 540
DB 498 CCAAGTAAATTTGAGAGTGAAGAAATGTTTAACTCAATGACTTGGACTT 557
QY 541 AGCTTCAAGAGGCAATCA-GGATTAAGTATATTCAGACAG-AACGGTATCCGGGCTC 598
DB 558 AGCTTCAAGAGGCAATCAAGATGTAAATGATTAAGAGAGACGGGTATCCGGGCTC 617
QY 599 CAGAGCAGATTTGCAAAATTTGCTGGCCAGAGGCTGGCTGAGAGTGA-TA-CAGATGT 657
DB 618 CAGAGCAGATTTGCAAAATTTGCTGGCCAGAGGCTGGCTGAGAGTGA-TA-CAGATGT 676
QY 658 ACCTCA 663
DB 677 ACCTCA 682

RESULT 13

BG386905

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

602454704P1 NIH_MGC_15 Homo sapiens CDNA clone IMAGE:4583065 5',
mRNA sequence.
BG386905.1 GI:13280454
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 906)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA library Preparation: Ling Hong/Rubin Laboratory
CDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: NIH Intramural Sequencing Center
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LHCN1306 row: 5 column: 02
High quality sequence stop: 668.
Location/Qualifiers
1..906
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4583065"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_11b="NIH MGC 15"
/note="Organ: colon; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dt priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGCAAGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)"

FEATURES

source

ORIGIN

Query Match 46.3%; Score 583.8; DB 4; Length 906;
Best Local Similarity 96.1%; Pred. No. 1.2e-159;
Matches 664; Conservative 0; Mismatches 17; Indels 10; Gaps 6;

QY 570 GTATATTCAGACAGCGGATTCGGGCTTCAGAGCAGAAATTCGAAATTCGTTGGCCCA 629
DB 2 GTATATTCAGACAGCGGATTCGGGCTTCAGAGCAGAAATTCGAAATTCGTTGGCCCA 61
QY 630 GGCTGGCTGCGAGTGTATACAGATGTATACCTCACTGTGATCTGTGAGAGCTGGAAT 689
DB 62 GGCTGGCTGCGAGTGTATACAGATGTATACCTCACTGTGATCTGTGAGAGCTGGAAT 121
QY 690 CATTATTCAGAAATGTTCTCAGAAATGAACTGAAACATACAGTCAAGTCTCAGAAATG 749
DB 122 CATTATTCAGAAATGTTCTCAGAAATGAACTGAAACATACAGTCAAGTCTCAGAAATG 181
QY 750 GAAGGCAACAGTTGCTTATTTATTCATCATATTTGCGAGTAAAGCAGTGTGAATGC 809
DB 182 GAAGGCAACAGTTGCTTATTTATTCATCATATTTGCGAGTAAAGCAGTGTGAATGC 241
QY 810 CGCAATTCAGAGCTTACCTTAAGAGACCTTTCAAAGCAGTCTTCAAGTCAAG 869
DB 242 CGCAATTCAGAGCTTACCTTAAGAGACCTTTCAAAGCAGTCTTCAAGTCAAG 301
QY 870 CAGAAAGATTCCTGCTGAATGAGCATTTGCGAGCCCATTCCTTACATTCCTTTGGCCC 929

Db 302 CAAAGAAATTCCTGCTGAATGCGATGTCAGCCCATCTTTCATTCCTTTGCCCC 361
 QY 930 TCATTTGAAGATCTGGTCAATGCTTCCACTCCAGTCTAAGATGCTGAATGCTGGA 989
 Db 362 TCATTTGAAGATCTGGTCAATGCTTCCACTCCAGTCTAAGATGCTGAATGCTGGA 421
 QY 990 TGATGATTTCTTGGGATGAGAGGAATGAGATGTTGTAAGATGTAAGAGGA 1049
 Db 422 TGATGATTTCTTGGGATGAGAGGAATGAGATGTTGTAAGATGTAAGAGGA 481
 QY 1050 GTGTCAAAAATATGACCAAGTGTCTACTTGTCCAAAGAAAATCTGCGAGAGG 1109
 Db 482 GTGTCAAAAATATGACCAAGTGTCTACTTGTCCAAAGAAAATCTGCG--AGAGG 539
 QY 1110 ACAAGCTTTGTGATGATGCAATGCTGATGATCCAAAGCTGCGGAGAAATTAAGTAC 1169
 Db 540 ACAAGCTTTGTGATGATGCAATG--TGGTGAATCCAAAGCTGCGGAGAAATTAAGTAC 598
 QY 1170 TGGAGAGATGTTGATGAGGAAGTTTGTGTGCTACATTCACCCGCTGAGTCTTACAA 1229
 Db 599 TGGAG--ATGTTGATGAGGAAGTTTGT--GGTACATTCACCCGCTGAGTCT--TACA 651
 QY 1230 GAGGGATATCTGTATCAAACTTCTCTTAA 1260
 Db 652 AGAGGATATCTGTATCAAACTGCGCTTAA 682

RESULT 14
 LOCUS BQ225071 888 bp mRNA linear EST 02-MAY-2002
 DEFINITION AGENCOURT 732333 NIH_MGC_70 Homo sapiens cDNA IMAGE:6016796
 5', mRNA sequence.

BQ225071
 BQ225071
 BQ225071.1 GI:20406471

EST.
 Homo sapiens (human)

SOURCE
 Homo sapiens

ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 NIH-MGC http://mgc.nci.nih.gov/
 Mammalian Gene Collection (MGC)

AUTHORS
 National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
 Unpublished (1999)

COMMENT
 Contact: Robert Strauberg, Ph.D.
 Email: cgaops-remail.nih.gov
 Tissue Procurement: ARCC

cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov

Plate: LLMK13215 row: a column: 21
 High quality sequence stop: 581.

FEATURES
 Location/Qualifiers

1. 888
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6016796"
 /tissue_type="epithelioid carcinoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_1ib="NIH MGC 70"
 /note="Organ: pancreas; Vector: pCMV-SPORTS; Site:1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.1 kb. Library constructed by Life
 Technologies."

ORIGIN

Query Match 46.0%; Score 579.6; DB 5; Length 888;
 Best Local Similarity 97.9%; Pident No. 2e-158;
 Matches 619; Conservative 0; Mismatches 9; Indels 4; Gaps 3;

QY 1 ATGCGCGGATCCCGCTGCGCCCTGCGGCGCGAGCCGCGCTTTCTGAGGCTTTGCGG 60
 Db 149 ATGCGCGGATCCCGCTGCGCCCTGCGGCGCGAGCCGCGCTTTCTGAGGCTTTGCGG 208
 QY 61 CGGCTGTGGAGGTACAGAGCGGCTCTGGGTACGGCTCTCGGCTCTGGGTATGCGGTT 120
 Db 209 CGGCTGTGGAGGTACAGAGCGGCTCTGGGTACGGCTCTCGGCTCTGGGTATGCGGTT 268
 QY 121 CGGCTGTGGAGGCAACCTGGCTCGCCCGCGGCGCTCAAGCAGTTCTTGCGCGAGGA 180
 Db 269 CGGCTGTGGAGGCAACCTGGCTCGCCCGCGGCGCTCAAGCAGTTCTTGCGCGAGGA 328
 QY 181 ACCACCGGGGCTCGGCTCTGCGCGCGGATATGTTTCGCAAGAGAGGCGGCGCTG 240
 Db 329 ACCACCGGGGCTCGGCTCTGCGCGCGGATATGTTTCGCAAGAGAGGCGGCGCTG 388
 QY 241 GAACAGTTGCGAGGTACAGAAACATCGTACTTTGTATGAGAGTCTTCAATCCACTTT 300
 Db 389 GAACAGTTGCGAGGTACAGAAACATCGTACTTTGTATGAGAGTCTTCAATCCACTTT 448
 QY 301 TCTCCAAATGTGCGATCAGCTGTCTGTGCTTGAATCTCTGAGATGCTGATTTGCGAA 360
 Db 449 TCTCCAAATGTGCGATCAGCTGTCTGTGCTTGAATCTCTGAGATGCTGATTTGCGAA 508
 QY 361 TTGCTTTATTTATTCAGTACAGAGGTTGTCATGTGATATACAGATTGCGCCGA 420
 Db 509 TTGCTTTATTTATTCAGTACAGAGGTTGTCATGTGATATACAGATTGCGCCGA 568
 QY 421 GATGTTTGGAGG--CCCTTGCTTTTCTTATCATAGAGGCTATGTCACGCGGACTCA 479
 Db 569 GATGTTTGGAGG--CCCTTGCTTTTCTTATCATAGAGGCTATGTCACGCGGACTCA 628
 QY 480 ACCACGTAACATATTTGTGAGTGCAGAGATGTAATTTTAACTCATGACTTTGAGACT 539
 Db 629 ACCACGTAACATATTTGTGAGTGCAGAGATGTAATTTTAACTCATGACTTTGAGACT 688
 QY 689 TACCTTCAAGAAAGGAAATCAAGATGTAAGTATTCAGACCGGCTTATGCGGCGC 748
 Db 597 TCCAGAGCAGAAATTCGAAATTTGCTTGCGCC 628
 QY 749 TCCAGAGCAGAAATTCGAAATTTGCTTGCGCC 780

RESULT 15
 LOCUS COJ37403 602 bp mRNA linear EST 29-JUL-2004

DEFINITION S1L702c17116f1 squirrel testis library 1 Spermophilus lateralis
 cDNA clone 17116 5', mRNA sequence.

COJ37403
 COJ37403
 COJ37403.1 GI:50824673

EST.
 Spermophilus lateralis (golden-mantled ground squirrel)

SOURCE
 Spermophilus lateralis

ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Scuriognathi; Scuridae; Scurinae;
 Spermophilus.

REFERENCE
 1 (bases 1 to 602)
 Williams,D.R., Gracey,A.Y., Martin,S.L., Hughes,M.A., Li,W.,
 Rogers,J. and Cossins,A.R.
 Microarray analysis of transcriptional changes during hibernation
 in the golden mantled ground squirrel, Spermophilus lateralis
 Unpublished (2004)

JOURNAL
 Contact: Andrew R. Cossins
 Laboratory for Environmental Gene Regulation
 University of Liverpool
 School of Biological Sciences, The Biosciences Building, Crown
 Street, Liverpool, United Kingdom, L69 7ZB
 Tel: +44(0)151-795-4510
 Fax: +44(0)151-795-4431
 Email: cossins@lly.ac.uk

COMMENT
 Contact: Andrew R. Cossins
 Laboratory for Environmental Gene Regulation
 University of Liverpool
 School of Biological Sciences, The Biosciences Building, Crown
 Street, Liverpool, United Kingdom, L69 7ZB
 Tel: +44(0)151-795-4510
 Fax: +44(0)151-795-4431
 Email: cossins@lly.ac.uk

Vector has been trimmed from this EST.

Plate: 17 row: i column: 16

Seq primer: pf1c T7 (5'-AATACGACTCACTATAGG-3')

High quality sequence stop: 602.

Location/Qualifiers

1..602

/organism="Spermophilus lateralis"

/mol_type="mRNA"

/db_xref="taxon:76772"

/clone="17116"

/sex="Male & female"

/tissue_type="Testis"

/dev_stage="Adult"

/lab_host="R. coli Electromax DH10B"

/note="Vector: pF1c; Site_1: Sail GTCAG; Site_2: BamHI

GATCC; Normalized and subtracted cDNA library prepared

from testis of hibernating and summer animals"

ORIGIN

Query Match 44.4%; Score 560; DB 7; Length 602;

Best Local Similarity 95.8%; Pred. No. 1e-152;

Matches 575; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

```
QY 624 GGGCCAGGCTGGCTGCAGAGTATACAGAAATGATACCTGCTGTGATCTGGAGCCT 683
DB 3 GGGCCAGGCTGGCTGCAGAGTATACAGAAATGATACCTGCTGTGATCTGGAGCCT 62
QY 684 AGGAATCAATTTTACTGGAATGTTCTCAGAAATGAACTGMAACATACAGTCAGATCTCA 743
DB 63 AGGAATTAATTTTACTGGAATGTTCTCAGAAATGAACTGMAACATACAGTCAGATCTCA 122
QY 744 GGAATGGAAGGCAACAGTCTCTATATTATGATCACTATTGGCCAGTAAAGCAGTGGT 803
DB 123 GGAATGGAAGGCAACAGTCTCTATATTATGATCACTATTGGCCAGTAAAGCAGTGGT 182
QY 804 GAATGCCGCAATCCAGCCTATACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGA 863
DB 183 GAATGCCGCAATCCAGCCTATACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGA 242
QY 864 TCCAGCAGAAAGATTCCTGCTGAAATGGCAATGTGCAAGCCATTCCTTAGCATTCCTTT 923
DB 243 TCCAGCAGAAAGATTCCTGCTGAAATGGCAATGTGCAAGCCATTCCTTAGCATTCCTTT 302
QY 924 TGGCCCTCATATTGAAGATGTGTCATGCTTCCCACTCCAGTCTAAGACTGCTGAATGT 983
DB 303 TGGCCCTCATATTGAAGATGTGTCATGCTTCCCACTCCAGTCTAAGACTGCTGAATGT 362
QY 984 GCTGATGATGATTAATCTTGGGAATGAAGAAATGAAATGATGTTGTAAGATGTAA 1043
DB 363 CTTGATGATGATTAATCTTGGGAATGAAGAAATGAAATGATGTTGTAAGATGTAA 422
QY 1044 AGAGGAGTGCAGAAAATATGACCAAGTGTATCTACTTGTCCAAAGAAAATCCTGG 1103
DB 423 AGAGGAGTGCAGAAAATATGACCAAGTGTATCTACTTGTCCAAAGAAAATCCTGG 482
QY 1104 CAGAGGCAAGTCTTGTGTGATGATGCAAAATGCTGGTATCCAAAGCTGCGCAAAAT 1163
DB 483 CAGAGGCAAGTCTTGTGTGATGATGCAAAATGCTGGTATCCAAAGCTGCGCAAAAT 542
QY 1164 ACTGACTGGAAGGATGTTTGTGAGGAAGTGTGTTGGCTACATTCACCCGCTGAGTGC 1223
DB 543 ACTGACTGGAAGGATGTTTGTGAGGAAGTGTGTTGGCTACATTCACCCGCTGAGTGC 602
```

Search completed: November 29, 2004, 17:46:23
Job time : 5316 secs

This Page Blank (uspto)


```
Db 181 ACCACGGGGCTCGGCTCTGCGCCGAGTATGTTTCGCAAGAGAGGGGGGCTG 240
Qy 241 GAACAGTTGAGGGTCAAGAAATCATGTAATTTGATGAGAGTGTAAATCCACTT 300
Db 241 GAACAGTTGAGGGTCAAGAAATCATGTAATTTGATGAGAGTGTAAATCCACTT 300
Qy 301 TCTCAAAATGTCATCAAGCTCTGTTGCTTGAATCTCTGATGTCAGTGTTCGAA 360
Db 301 TCTCAAAATGTCATCAAGCTCTGTTGCTTGAATCTCTGATGTCAGTGTTCGAA 360
Qy 361 TTGCTCTTAATTCAGTCAAGGGTGTTCATGTAATGATGATGATGATGATGATG 420
Db 361 TTGCTCTTAATTCAGTCAAGGGTGTTCATGTAATGATGATGATGATGATGATG 420
Qy 421 GATGTTTGAAGGCTGCTGTTTCTTCATCAAGAGGCTATGATGATGATGATGATG 480
Db 421 GATGTTTGAAGGCTGCTGTTTCTTCATCAAGAGGCTATGATGATGATGATGATG 480
Qy 481 CCACGTAAATATGTAAGTGAAGTGAATGATGATGATGATGATGATGATGATGATG 540
Db 481 CCACGTAAATATGTAAGTGAAGTGAATGATGATGATGATGATGATGATGATGATG 540
Qy 541 AGCTTCAAAAGAGGCAATCAGATGTAAGTATTAATCAAGACGAGTATCGGGCTCA 600
Db 541 AGCTTCAAAAGAGGCAATCAGATGTAAGTATTAATCAAGACGAGTATCGGGCTCA 600
Qy 601 GAAGGAGATTTGCAAAATGCTTGGCCAGGCTGCGCTGAGATGATCAGAAATGATC 660
Db 601 GAAGGAGATTTGCAAAATGCTTGGCCAGGCTGCGCTGAGATGATCAGAAATGATC 660
Qy 661 TCAGCTGTGATCTGTGAGGCTTGAATCATTTTACTGGAATGTTTCTCAGGAATGAA 720
Db 661 TCAGCTGTGATCTGTGAGGCTTGAATCATTTTACTGGAATGTTTCTCAGGAATGAA 720
Qy 721 CTGAAACATACAGTCAATCTCAGAAATGGAAGCAACAGTCTGTATTAATGATCAC 780
Db 721 CTGAAACATACAGTCAATCTCAGAAATGGAAGCAACAGTCTGTATTAATGATCAC 780
Qy 781 ATATTGCGCAGTAAAGAGTGAATGCGCAATTCAGCTATACCTAAGAGACCTT 840
Db 781 ATATTGCGCAGTAAAGAGTGAATGCGCAATTCAGCTATACCTAAGAGACCTT 840
Qy 841 ATCAAAAGCATGCTTATGATGATCCAAAGCAAGAAATTCCTGCTGAATGCGATGTGC 900
Db 841 ATCAAAAGCATGCTTATGATGATCCAAAGCAAGAAATTCCTGCTGAATGCGATGTGC 900
Qy 901 AGCCCATCTTTAGCATTCCTTTTGGCCCTCATATTTGAAGATCGTCAATGCTTCCACT 960
Db 901 AGCCCATCTTTAGCATTCCTTTTGGCCCTCATATTTGAAGATCGTCAATGCTTCCACT 960
Qy 961 CCAGTCTAAGATGCTGAATGCTGAATGATGATGATGATGATGATGATGATGATG 1020
Db 961 CCAGTCTAAGATGCTGAATGCTGAATGATGATGATGATGATGATGATGATGATG 1020
Qy 1021 GAAGATGTTGTAAGATGTAAGAGAGATGTAAGAAATATGAAATGCAAGTGTATCTTA 1080
Db 1021 GAAGATGTTGTAAGATGTAAGAGAGATGTAAGAAATATGAAATGCAAGTGTATCTTA 1080
Qy 1081 CTGTGTTCAAAAGAAATCTGCGCAGAGCAAGTCTTTGTTGATGTAATGCTGTGT 1140
Db 1081 CTGTGTTCAAAAGAAATCTGCGCAGAGCAAGTCTTTGTTGATGTAATGCTGTGT 1140
Qy 1141 GATTCCAAAGCTGCGCAGAAATTAATGATGTAAGATGTTTGAAGAGTTGTTGTG 1200
Db 1141 GATTCCAAAGCTGCGCAGAAATTAATGATGTAAGATGTTTGAAGAGTTGTTGTG 1200
Qy 1201 GCTACATTTACCCGCTGAGTGCCTACAGAGGGGATATCTGTATCAAACTTGTCTTA 1260
Db 1201 GCTACATTTACCCGCTGAGTGCCTACAGAGGGGATATCTGTATCAAACTTGTCTTA 1260
```

RESULT 2

```
AX262521
LOCUS AX262521 1260 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 12 from Patent WO0173050.
ACCESSION AX262521
VERSION AX262521.1 GI:16511408
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1. Meyers, R. 3714, 16742, 23546, and 13887 novel protein kinase molecules and
    uses therefor Patent: WO 0173050-A 12 04-OCT-2001;
    Millennium Pharmaceuticals, Inc. (US)
FEATURES
    source
        location/Qualifiers
            1..1260
                /organism="Homo sapiens"
                /mol_type="unassigned DNA"
                /db_xref="taxon:9606"
ORIGIN
Query Match 99.7%; Score 1256.8; DB 6; Length 1260;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1258; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGGCGGGATCCGGCTGCGCTGCGGCGCGAGCCGCGCTTTCTGAGGCTTCCGG 60
Db 1 ATGGCGGGATCCGGCTGCGCTGCGGCGCGAGCCGCGCTTTCTGAGGCTTCCGG 60
Qy 61 CCGCTGTGCAAGTACAGAGCCGTCTGGGTAGCCGCTCTCCGCTCGGTATCGGTT 120
Db 61 CCGCTGTGCAAGTACAGAGCCGTCTGGGTAGCCGCTCTCCGCTCGGTATCGGTT 120
Qy 121 CCGTGTGCGCAACCTTGCTGCGCCCGCGGCGCTTCAAGAGTCTTGGCCGAGAA 180
Db 121 CCGTGTGCGCAACCTTGCTGCGCCCGCGGCGCTTCAAGAGTCTTGGCCGAGAA 180
Qy 181 ACCACGGGGCTGCGGCTCTGCTGCGCGAGTATGTTTCCGAAAGAGAGGCGGCTG 240
Db 181 ACCACGGGGCTGCGGCTCTGCTGCGCGAGTATGTTTCCGAAAGAGAGGCGGCTG 240
Qy 241 GAACAGTTGCAAGGCTCACAGAAATCGTGAATTTGATGAGTGTAAATCAATCACTT 300
Db 241 GAACAGTTGCAAGGCTCACAGAAATCGTGAATTTGATGAGTGTAAATCAATCACTT 300
Qy 301 TCTCCAAATGTCATCAAGCTCTGTTGCTTGAATCTCTGATGTCAGTGTTCGAA 360
Db 301 TCTCCAAATGTCATCAAGCTCTGTTGCTTGAATCTCTGATGTCAGTGTTCGAA 360
Qy 361 TTGCTCTTAATTCAGTCAAGGGTGTTCATGTAATGATGATGATGATGATGATGATG 420
Db 361 TTGCTCTTAATTCAGTCAAGGGTGTTCATGTAATGATGATGATGATGATGATGATG 420
Qy 421 GATGTTTGAAGGCTTGTCTTCTTCATCAAGAGGCTATGATGATGATGATGATGATG 480
Db 421 GATGTTTGAAGGCTTGTCTTCTTCATCAAGAGGCTATGATGATGATGATGATGATG 480
Qy 481 CCACGTAAATATGTAAGTGAAGTGAATGATGATGATGATGATGATGATGATGATG 540
Db 481 CCACGTAAATATGTAAGTGAAGTGAATGATGATGATGATGATGATGATGATGATG 540
Qy 541 AGCTTCAAAAGAGCAATCAGATGTAAGTATTAATCAAGACGAGGATCGGCTCCA 600
Db 541 AGCTTCAAAAGAGCAATCAGATGTAAGTATTAATCAAGACGAGGATCGGCTCCA 600
Qy 601 GAAGGAGATTTGCAAAATGCTTGGCCAGGCTGCGCTGAGATGATCAGAAATGATC 660
Db 601 GAAGGAGATTTGCAAAATGCTTGGCCAGGCTGCGCTGAGATGATCAGAAATGATC 660
Qy 661 TCAGCTGTGATCTGTGAGGCTAAGAAATCATTTTATGTAATGTTTCTCAGGAATGAA 720
```

```

Db      TCAGCTGTGATCTGTGAGCCTAGGAATCATTTTACTGGAATGTTCTCAGGAATGAA 720
Qy      CTGAAACATACGTGATCTCAGGAATGGAAGGCAACAGTTCTGCTATTATGATCAC 780
Db      CTGAAACATACGTGATCTCAGGAATGGAAGGCAACAGTTCTGCTATTATGATCAC 780
Qy      ATATTTGCGAGAAAGAGTGTGATGCGCAATTCAGCCTTATCACTTAAGAGACTT 840
Db      ATATTTGCGAGAAAGAGTGTGATGCGCAATTCAGCCTTATCACTTAAGAGACTT 840
Qy      ATCAAAAGCATGCTTCATGATGATCCAGCAGAAAGATTCTGCTGAAATGCGATTGTC 900
Db      ATCAAAAGCATGCTTCATGATGATCCAGCAGAAAGATTCTGCTGAAATGCGATTGTC 900
Qy      AGCCCATTTTGAAGATTCCTTTTGGCCCTCATATTGAAGATCTGCTATGCTCCACT 960
Db      AGCCCATTTTGAAGATTCCTTTTGGCCCTCATATTGAAGATCTGCTATGCTCCACT 960
Qy      CCAGTGTCTAAGATGCTGAAATGCTGATGATGATATCTTGGGAATGGAAGGAATAT 1020
Db      CCAGTGTCTAAGATGCTGAAATGCTGATGATGATATCTTGGGAATGGAAGGAATAT 1020
Qy      GAAAGATGTTGTAAGATGTAAAGAGAGTGTCAAAAATATGACCAAGTGTATCTCTA 1080
Db      GAAAGATGTTGTAAGATGTAAAGAGAGTGTCAAAAATATGACCAAGTGTATCTCTA 1080
Qy      CTTGTTCCAAAGAAATCTCTGCGCAGAGACAAAGTCTTTGTTGATGATGCAATCTGCT 1140
Db      CTTGTTCCAAAGAAATCTCTGCGCAGAGACAAAGTCTTTGTTGATGATGCAATCTGCT 1140
Qy      GATTCCAAAGTGTGCGCAGAAATTAAGTCTGAGATGATGATGATGATGATGATGATG 1200
Db      GATTCCAAAGTGTGCGCAGAAATTAAGTCTGAGATGATGATGATGATGATGATGATG 1200
Qy      GCTACATTTCTACCCGCTGAGTGTCTTCAAGAGGAGATCTGTATCAAACTTGTCTTAA 1260
Db      GCTACATTTCTACCCGCTGAGTGTCTTCAAGAGGAGATCTGTATCAAACTTGTCTTAA 1260

RESULT 3
AX262519      2598 bp      DNA      linear      PAT 26-OCT-2001
LOCUS      AX262519
DEFINITION      Sequence 10 from Patent WO0173050.
ACCESSION      AX262519
VERSION      AX262519.1 GI:16511406
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1      Meyers, R.
AUTHORS      3714, 16742, 23546, and 13887 novel protein kinase molecules and
TITLES      uses therefor
JOURNAL      Patent: WO 0173050-A 10 04-OCT-2001;
MILLENNIUM Pharmaceuticals, Inc. (US)
FEATURES
Source
1. .2598
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
268..1527
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD10238.1"
/db_xref="GI:16511407"
/translation="MAGSGCAGWAGAPPRFLAFLRLMQVSHLSSSSASVYRVACCG
NGSPGALKQFLPGTTGAASAABVGRKRALBOLQGRNVTLYGVTFHSP
NPSRCLILELDVGSVSLYSHOGMMWIOHCAADYVLAFLHREGVHADIK
PNVITMSAENECFKLIDFGLSPKGNQVYKTYQTDGYAPRAELQNCALQAGLSDTE
CTSAVDLWSLGIILLEMFGMLKHTVSKQEMKANSSAIDHIFASKAVVNAISAYH
LADLTKSMHDDPSRRIPALMAAGSPFESIPAPRIEDLVMLPTVLRLNLVLDLYL
ENBEYEDVDEKESQKYGVVSLVAKENPRGQVFEVYANAGDSKAAQKLTLGR

```

```

ORIGIN      MFQDKFVATFYPPLSAVKRYLYQTLI"
Query Match      99.7%; Score 1256.8; DB 6; Length 2598;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1258; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy      1      ATGGCGGAGATCCGGCTGCGCTGGGGGCGGAGCGCGCGCTTTCTGAGAGCTTCGGG 60
Db      ATGGCGGAGATCCGGCTGCGCTGGGGGCGGAGCGCGCGCTTTCTGAGAGCTTCGGG 327
Qy      61      CGGCTGTGGCAGATACAGAGCCGTCTGGGTAGCGGCTCTCCGCTCGGTGTATCGGGTT 120
Db      CGGCTGTGGCAGATACAGAGCCGTCTGGGTAGCGGCTCTCCGCTCGGTGTATCGGGTT 387
Qy      121     CGCTGTGGCGAACCCTGGCTGCGCCCGCGCGCTCTCAAGCAGTTCTTGCCGCGCAGA 180
Db      CGCTGTGGCGAACCCTGGCTGCGCCCGCGCGCTCTCAAGCAGTTCTTGCCGCGCAGA 447
Qy      181     ACCACCGGGGCTGGGCGCTCTGCGCGAGTATGGTTCCGCAAGAGAGGGCGGCGCTG 240
Db      ACCACCGGGGCTGGGCGCTCTGCGCGAGTATGGTTCCGCAAGAGAGGGCGGCGCTG 507
Qy      241     GAAACAGTTGCAAGGTCACAGAAACATCTGTGATTGTATGAGAGTGTTCACATCCATT 300
Db      GAAACAGTTGCAAGGTCACAGAAACATCTGTGATTGTATGAGAGTGTTCACATCCATT 567
Qy      301     TCTCCAAATGTCACATCAGCGCTGTCTGTGCTTGAATCTCTGATGTGATGTTGGAA 360
Db      TCTCCAAATGTCACATCAGCGCTGTCTGTGCTTGAATCTCTGATGTGATGTTGGAA 627
Qy      361     TTGCTCTTATTTTCAGATCACAGGGTGTCCAATGAGATGATACAGATTCGCGCCGA 420
Db      TTGCTCTTATTTTCAGATCACAGGGTGTCCAATGAGATGATACAGATTCGCGCCGA 687
Qy      421     GATGTTTGGAGGCGCTTCTTCTTCATCATGAGGGCTATGTCATGCGGACCTCAA 480
Db      GATGTTTGGAGGCGCTTCTTCTTCATCATGAGGGCTATGTCATGCGGACCTCAA 747
Qy      481     CCACGTAAATATTTGTGAGATGTCAGAGATGATGTTTAAATCTCATTTGACTTGGACT 540
Db      CCACGTAAATATTTGTGAGATGTCAGAGATGATGTTTAAATCTCATTTGACTTGGACT 807
Qy      541     AGCTTCAAGAAAGGAATCAGAGATGTAAGTATTTCAAGACAGGCTATCGGCTCCA 600
Db      AGCTTCAAGAAAGGAATCAGAGATGTAAGTATTTCAAGACAGGCTATCGGCTCCA 867
Qy      601     GAAGCAGAAATTCAGAAATTTGCTGGCCAGGCTGCGCTGCAAGTATACAGAAATGACC 660
Db      GAAGCAGAAATTCAGAAATTTGCTGGCCAGGCTGCGCTGCAAGTATACAGAAATGACC 927
Qy      661     TCAGCTGTGATCTGTGAGCCTTAGAATCATTTTACTGGAATGTTCTCAGGAATGAA 720
Db      TCAGCTGTGATCTGTGAGCCTTAGAATCATTTTACTGGAATGTTCTCAGGAATGAA 987
Qy      721     CTGAAACATACGTGATCTCAGGAATGGAAGGCAACAGTTCTGCTATTATGATCAC 780
Db      CTGAAACATACGTGATCTCAGGAATGGAAGGCAACAGTTCTGCTATTATGATCAC 1047
Qy      781     ATATTTGCGAGTAAAGAGTGTGATGCGCAATTCAGCCTTATCACTTAAGAGACTT 840
Db      ATATTTGCGAGTAAAGAGTGTGATGCGCAATTCAGCCTTATCACTTAAGAGACTT 1107
Qy      841     ATCAAAAGCATGCTTCATGATGATCCAGCAGAAAGATTCTGCTGAAATGCGATTGTC 900
Db      ATCAAAAGCATGCTTCATGATGATCCAGCAGAAAGATTCTGCTGAAATGCGATTGTC 1167
Qy      901     AGCCCATTTTGAAGATTCCTTTTGGCCCTCATATTGAAGATCTGCTATGCTCCACT 960
Db      AGCCCATTTTGAAGATTCCTTTTGGCCCTCATATTGAAGATCTGCTATGCTCCACT 1227
Qy      961     CCAGTGTCTAAGATGCTGAAATGCTGATGATGATATCTTGGGAATGGAAGGAATAT 1020

```

Db	1228	CGAGTGTAAAGCTGCTGAATGTCGTGAGTGAATGATTAATCTTGAGAAATGAAGAGAAATAT	1287
Qy	1021	GAAGATGTTGTGAAGAATGTAAAGAAGAGAGTGTCAAAAATATGACACCAAGTGTATCTCTA	1080
Db	1288	GAAGATGTTGTGAAGAATGTAAAGAAGAGAGTGTCAAAAATATGACACCAAGTGTATCTCTA	1347
Qy	1081	CTTGTTCCAAAGGAAATCTCTGACAGAGACAAGTCTTTGTTGAGTATGCAATGCTGAT	1140
Db	1348	CTTGTTCCAAAGGAAATCTCTGACAGAGACAAGTCTTTGTTGAGTATGCAATGCTGAT	1407
Qy	1141	GATTCCAAAGCTGGCGCAAAATTAATGACTGGAAGATGTTTGAGTGGAACTTTGTTGTG	1200
Db	1408	GATTCCAAAGCTGGCGCAAAATTAATGACTGGAAGATGTTTGAGTGGAACTTTGTTGTG	1467
Qy	1201	GCTACATTTCAACCGCTGAGAGCTCTACAAAGAGGAAATCTGATCAAACTTGGCTTAA	1260
Db	1468	GCTACATTTCAACCGCTGAGAGCTCTACAAAGAGGAAATCTGATCAAACTTGGCTTAA	1527
RESULT 4			
HSAS36197			
LOCUS	HSAS36197	2909 bp	mRNA
DEFINITION	Homo sapiens mRNA for KIS protein.	linear	PR1 11-JAN-2003
ACCESSION	AJ536197		
VERSION	AJ536197.1	GI:27657360	
KEYWORDS	KIS protein.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eulheia; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	Bleche, I., Manceau, V., Curmi, P. A., Laurendon, I., Iachkar, S., Leroy, K., Vaidan, D., Sobel, A. and Maucuer, A.		
TITLE	Quantitative RT-PCR reveals a ubiquitous but preferentially neural expression of the KIS gene in rat and human		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2909)		
AUTHORS	Maucuer, A.		
TITLE	Direct Submission		
JOURNAL	Submitted (08-JAN-2003) Maucuer A., U440, Inserm, IFM, 17 rue du fer a moulin, 75005 Paris, FRANCE		
FEATURES	Location/Qualifiers		
source	1..2909		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="RZPD:IMAG0998W213680"		
	/db_xref="taxon:9606"		
	/clone="I.M.A.G.E. Consortium CloneID 1455548"		
	/rfeature type="2 pooled tumors (clear cell type)"		
gene	1..2909		
	/gene="KIS"		
CDS	146..1405		
	/gene="KIS"		
	/codon_start=1		
	/product="KIS protein"		
	/protein_id="CAD60192.1"		
	/db_xref="GI:27657361"		
	/translation="MAGSGCAGMGAEPRLLEAFRLMVOYSLGSSASRYRRCCG NPSPEGLAKQFLPFGTGTAAASAAYEGRKRAALDEQLQGRNIVTIGVFTTHFSP NVPSRCLLEELDVSVSELVLSHQGCSMMMIQHCARDVLEALAFLHREGVVAHDE PKRIIVLSAENECCKFLIDFGLSFEKGNQDVXYIQDGYRAPELQNCIAQGLSDTE CTGAVDMSLGIILLEMFSQMKLHTFRSIOEWKANSASAIIDHI FASKVAVNAIDAYH LRLDLKMLHDDPSRRI PALMALCSPFSI PRAPHIEDLVMLPTVLRLLNVLDQDYL ENEEBVDVDEYDEKCCOKGKPYSLVLEKENGQGVVEVYANAGDSKAQKLITGR MFGKRVVATFYFLSAIKRGYLYQTLL"		
ORIGIN			
Query Match	99.7%; Score 1256.8; DB 9; Length 2909;		
Best Local Similarity	99.8%; Pred. No. 0;		
Matches 1258; Conservative	0; Mismatches 2; Indels 0; Gaps 0;		
Qy	1	ATGGCGGGATTCGGGCTGCGCTGGAGCGCGAGCCGCGGCTTTTCTGAGGCTTTGGG	60
Db			
Db	146	ATGGCGGAAATCCGGCTGCGCTGGAGCGCGAGCCGCGGCTTTTCTGAGGCTTTGGG	205
Qy	61	CGGCTGTGCAAGTACAGAGCCGCTCTGGGTATAGCGGCTCTCCGCTCGGTTATCGGGT	120
Db	206	CGGCTGTGCAAGTACAGAGCCGCTCTGGGTATAGCGGCTCTCCGCTCGGTTATCGGGT	265
Qy	121	CGGCTGTGCGGACACCCCTGCTCGGCCCGCGGCGGCTCTCAAGCACTTCTTCCGCA	180
Db	266	CGGCTGTGCGGGAACCTTGCTCGGCCCGCGGCGGCTCTCAAGCACTTCTTCCGCA	325
Qy	181	ACCAACGGGGCTGCGGCTCTGCGGCCGAGATAGGTTTCCGCAAGAGAGGGCGGCTG	240
Db	326	ACCAACGGGGCTGCGGCTCTGCGGCCGAGATAGGTTTCCGCAAGAGAGGGCGGCTG	385
Qy	241	GAACAGTGCAGGGTACAGAAACATACGATCTTGTATGAGAGTGTATCAATCACTTT	300
Db	386	GAACAGTGCAGGGTACAGAAACATCGATCTTGTATGAGAGTGTATCAATCACTTT	445
Qy	301	TCTCCAAATGTGCATACGCTGTCTGTTCCTTGAATCTCGATGTGATGTTCCGAA	360
Db	446	TCTCCAAATGTGCATACGCTGTCTGTTCCTTGAATCTCGATGTGATGTTCCGAA	505
Qy	361	TGCTCTTATATTTCCAGTACACAGGGTGTTCACATGTAGATATACAGATGAGCCGA	420
Db	506	TGCTCTTATATTTCCAGTACACAGGGTGTTCACATGTAGATATACAGATGAGCCGA	565
Qy	421	GATGTTTGGAGGCCCTTGCTTTTCTTATCATCATGAGGCTATGTCCATCGGACTCAA	480
Db	566	GATGTTTGGAGGCCCTTGCTTTTCTTATCATCATGAGGCTATGTCCATCGGACTCAA	625
Qy	481	CCACGTATCATTTGTGAGTGCAGAGAAATGAATGTTTAAACATTAATGACTTGGACTT	540
Db	626	CCACGTATCATTTGTGAGTGCAGAGAAATGAATGTTTAAACATTAATGACTTGGACTT	685
Qy	541	AGCTTCAAGAGGCAATCAGATGTAAAGTATTTATTCAGACAGACGGGTATCGGGCTCA	600
Db	686	AGCTTCAAGAGGCAATCAGATGTAAAGTATTTATTCAGACAGACGGGTATCGGGCTCA	745
Qy	601	GAAGCAAGATTGCAAAATTCCTTGGCCAGAGCTGGCTGCGAGTGTATCAGATGTATCC	660
Db	746	GAAGCAAGATTGCAAAATTCCTTGGCCAGAGCTGGCTGCGAGTGTATCAGATGTATCC	805
Qy	661	TCAGCTGTGATCTGTGAGGCTCTAGGAATCATTTTACTGGAATGTTTCTCAGATGAAA	720
Db	806	TCAGCTGTGATCTGTGAGGCTCTAGGAATCATTTTACTGGAATGTTTCTCAGATGAAA	865
Qy	721	CTGAAACATACAGTCAAGTCTCAGGAATGGAAGGCAACAGTTGCTGATTTATTTGATCAC	780
Db	866	CTGAAACATACAGTCAAGTCTCAGGAATGGAAGGCAACAGTTGCTGATTTATTTGATCAC	925
Qy	781	ATATTGTCAGATTAAGCAATGAGTGAATGCGGAATTTCCAGCTATACCTTAAGAGACCTT	840
Db	926	ATATTGTCAGATTAAGCAATGAGTGAATGCGGAATTTCCAGCTATACCTTAAGAGACCTT	985
Qy	841	ATCAAAAGATGCTTCATGATGATCCAAGCAGAAAGATTCCTGCTGAATGGCAATTGTC	900
Db	986	ATCAAAAGATGCTTCATGATGATCCAAGCAGAAAGATTCCTGCTGAATGGCAATTGTC	1045
Qy	901	AGCCCAATCTTTAGCATTCCTTTTGGCCCTCATATTGAAGATCTGTATGCTTCCACT	960
Db	1046	AGCCCAATCTTTAGCATTCCTTTTGGCCCTCATATTGAAGATCTGTATGCTTCCACT	1105
Qy	961	CCAGTCTTAAGACTGCTGAATGAGTGTGATGATTTCTTGGGAATGAAGAGAAATAT	1020
Db	1106	CCAGTCTTAAGACTGCTGAATGAGTGTGATGATTTCTTGGGAATGAAGAGAAATAT	1165
Qy	1021	GAAAGATGTTGTGAAGATTAAGAAGAGAGTGTCAAAAATATGACACAGTGTATCTCTA	1080
Db	1166	GAAAGATGTTGTGAAGATTAAGAAGAGAGTGTCAAAAATATGACACAGTGTATCTCTA	1225
Qy	1081	CTTGTTCCAAAGGAAATCTCTGACAGAGACAAGTCTTTGTTGAGTATGCAATGCTGAT	1140
Db	1226	CTTGTTCCAAAGGAAATCTCTGACAGAGACAAGTCTTTGTTGAGTATGCAATGCTGAT	1285

Qy	1141	GATTCGCAAGCTGGCGGCACAAATTTATGACCTGACGAGAGAGTTGTATGATGGGAAGCTTGTGTG	1200
Db	1286	GATTCGCAAGCTGGCGGCACAAATTTATGACCTGACGAGAGAGTTGTATGATGGGAAGCTTGTGTG	1345
Qy	1201	GCTACATTTCTACCCGCTGAGTGCCTACAGAGGGGATATCTGTATCAAACTTGCTTTAA	1260
Db	1346	GCTACATTTCTACCCGCTGAGTGCCTACAGAGGGGATATCTGTATCAAACTTGCTTTAA	1405
RESULT 5			
LOCUS	CQ412057	4065 bp	DNA
DEFINITION	Sequence 19128 from Patent WO0170979.	linear	PAT 23-JAN-2004
ACCESSION	CQ412057		
VERSION	CQ412057.1	GI:41319838	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	Lee, J. and Lillie, J.		
JOURNAL	Genes, compositions, kits, and method for identification, assessment, prevention, and therapy of ovarian cancer		
FEATURES	Patent: WO 0170979-A 19128 27-SEP-2001; Millennium Pharmaceuticals, Inc. (US)		
source	Location/Qualifiers		
	1..4065		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match 99.7%; Score 1256.8; DB 6; Length 4065;			
Best Local Similarity 99.8%; Pred. No. 0;			
Matches 1258; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
Qy	1	ATGGCGGGATCCGCGTGGCGCTGGCGGGCGCGAGCGCGCGCTTCTGAGAGCCCTTGGG	60
Db	268	ATGGCGGGATCCGCGTGGCGCTGGCGGGCGCGAGCGCGCGCTTCTGAGAGCCCTTGGG	327
Qy	61	CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCTCCGCTCGGTATCGGTT	120
Db	328	CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCTCTCCGCTCGGTATCGGTT	387
Qy	121	CGCTGCTCGCGAACCTCTGGCTCGCCCGCGCGCTCTCAAGACAGTTCTTGCGCGACGA	180
Db	388	CGCTGCTCGCGAACCTCTGGCTCGCCCGCGCGCTCTCAAGACAGTTCTTGCGCGACGA	447
Qy	181	ACACCGGGGCGCGGCGCTCTGCGCGCGAGTAGTGTTCGCGAAGAGAGGGCGCGCTG	240
Db	448	ACACCGGGGCGCGGCGCTCTGCGCGCGAGTAGTGTTCGCGAAGAGAGGGCGCGCTG	507
Qy	241	GAACAGTTGACAGGTCACAGAAACATCGTGACTTTGTATGAGAGTGTTACATCACTTT	300
Db	508	GAACAGTTGACAGGTCACAGAAACATCGTGACTTTGTATGAGAGTGTTACATCACTTT	567
Qy	301	TCTCCAAATGTCACATCAGCGTGTCTGTGCTTGAACCTCTGATGTCAAGTGTGGAA	360
Db	568	TCTCCAAATGTCACATCAGCGTGTCTGTGCTTGAACCTCTGATGTCAAGTGTGGAA	627
Qy	361	TTGCTCTATATTCAGATCACAGAGGTTGTTCCATGTGATGATACAGATTTGCGCCGA	420
Db	628	TTGCTCTATATTCAGATCACAGAGGTTGTTCCATGTGATGATACAGATTTGCGCCGA	687
Qy	421	GATGTTTGGAGCCCTTGCTTTTCTTCAATCAATGAGGCTATGTCATGCGGACTCAAA	480
Db	688	GATGTTTGGAGCCCTTGCTTTTCTTCAATCAATGAGGCTATGTCATGCGGACTCAAA	747
Qy	481	CCAGCTACATTTGTGTGAGTGCACAGATGAATGTTTTAACTATTTGACTTTGACTT	540
Db	748	CCAGCTACATTTGTGTGAGTGCACAGATGAATGTTTTAACTATTTGACTTTGACTT	807

Qy	541	AGCTTCAAAGAGGCATCATGAGATGTAAAGTATATTTTCAGACAGACGGGTATCCGGCTCCCA	600
Db	808	AGCTTCAAAGAGGCATCATGAGATGTAAAGTATATTTTCAGACAGACGGGTATCCGGCTCCCA	867
Qy	601	GAAGCAGAAATTTGGCAAAATTTGCTTGCCCGCCAGGCTGCGCTGCGAGGTGATACAGAAATGTAAC	660
Db	868	GAAGCAGAAATTTGGCAAAATTTGCTTGCCCGCCAGGCTGCGCTGCGAGGTGATACAGAAATGTAAC	927
Qy	661	TCAGCTGTGATCTGTGAGACCTTAGAATCATTTTACTGGAATGTCTCAGGAATGTAAA	720
Db	928	TCAGCTGTGATCTGTGAGACCTTAGAATCATTTTACTGGAATGTCTCAGGAATGTAAA	987
Qy	721	CTGAAACATTCAGTCAGATCTCAGGAATGGAAGCAAAACAGTTCTGCTATTTATTTGATCAC	780
Db	988	CTGAAACATTCAGTCAGATCTCAGGAATGGAAGCAAAACAGTTCTGCTATTTATTTGATCAC	1047
Qy	781	ATATTTTCCAGTAAAGCAGTGGTGAATGCCGCAATTCAGACCTATCACTTAAGAGACTT	840
Db	1048	ATATTTTCCAGTAAAGCAGTGGTGAATGCCGCAATTCAGACCTATCACTTAAGAGACTT	1107
Qy	841	ATCAAAGCATGCTTTCATGATGATCCAGAGAGAAAGAAATTCCTGCTGAAATGGCAATTTGTGC	900
Db	1108	ATCAAAGCATGCTTTCATGATGATCCAGAGAGAAAGAAATTCCTGCTGAAATGGCAATTTGTGC	1167
Qy	901	AGCCCATTCCTTTAGCATTCCTTTTGGCCCTCATATTTGAAGATCTGTCATGCTTCCACT	960
Db	1168	AGCCCATTCCTTTAGCATTCCTTTTGGCCCTCATATTTGAAGATCTGTCATGCTTCCACT	1227
Qy	961	CCAGTGTCTAAGACTGCTGAAATGTGCTGGATGATGATATTAATCTTGGAATGGAAGAAATAT	1020
Db	1228	CCAGTGTCTAAGACTGCTGAAATGTGCTGGATGATGATATTAATCTTGGAATGGAAGAAATAT	1287
Qy	1021	GAAGATGTTGTGAAGAATGTAAAAAGAGACTGTCAAAATATATGGAACGAGTATCTCTA	1080
Db	1288	GAAGATGTTGTGAAGAATGTAAAAAGAGACTGTCAAAATATATGGAACGAGTATCTCTA	1347
Qy	1081	CTTGTTCCAAAGAAATATCTGCGACAGAGACAAGTCTTTGTGAGTATGCAAAATCTGTGT	1140
Db	1348	CTTGTTCCAAAGAAATATCTGCGACAGAGACAAGTCTTTGTGAGTATGCAAAATCTGTGT	1407
Qy	1141	GATTCCAAAGCTGCGAGAAATTAATCTGACTGGAAGGATGTTTGAATGGAAGATTTGTTGTG	1200
Db	1408	GATTCCAAAGCTGCGAGAAATTAATCTGACTGGAAGGATGTTTGAATGGAAGATTTGTTGTG	1467
Qy	1201	GCTACATTTTACCCGCTGAGATGCTTAATAAGGGGATATCTGATCAACCTTGCTTAA	1260
Db	1468	GCTACATTTTACCCGCTGAGATGCTTAATAAGGGGATATCTGATCAACACCTTGCTTAA	1527
RESULT 6			
AX188296	LOCUS	4065 bp	DNA
DEFINITION	Sequence 3991 from Patent WO0142467.	linear	PAT 06-AUG-2001
AX188296	VERSION		
AX188296.1	GI:15139769		
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
JOURNAL	Schlegel, R., Deeds, J., Berger, A. and Zhao, X.		
FEATURES	Genes, compositions, kits, and methods for identification,		
source	assessment, prevention, and therapy of cervical cancer		
	Patent: WO 0142467-A 3991.14-UN-2001;		
	Millennium Predictive Medicine, Inc. (US)		
	Location/Qualifiers		
	1. .4065		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			

Query Match 99.7%; Score 1256.8; DB 6; Length 4065;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1258; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 ATGGCGGATCCGGCTGCGCTGCGGCGCGGAGCGCGCGGTTTCTGAGGCTTCGCGG 60
DB ATGGCGGATCCGGCTGCGCTGCGGCGCGGAGCGCGCGGTTTCTGAGGCTTCGCGG 327
QY 61 CGGCTGTGCGAGGTACAGAGCGCTGTGGGTAGCGGCTCCCGCGCTCGGTATCGGGTT 120
DB 328 CGGCTGTGCGAGGTACAGAGCGCTGTGGGTAGCGGCTCCCGCGCTCGGTATCGGGTT 387
QY 121 CGCTGTGCGGCAACCGTGGCTGCGCCCGCGCGCGCTCAAGCAATTCCTTGCAGGCA 180
DB 388 CGCTGTGCGGCAACCGTGGCTGCGCCCGCGCGCGCTCAAGCAATTCCTTGCAGGCA 447
QY 181 ACCACCGGCGCTGCGGCTCTGCGCGCGAGTATGTTTCCGAAAGAGAGGCGCGCTG 240
DB 448 ACCACCGGCGCTGCGGCTCTGCGCGCGAGTATGTTTCCGAAAGAGAGGCGCGCTG 507
QY 241 GAACGTTGCGAGGTACAGAAACATCGTGAATTTGATGAGTGTTCATCCACTTT 300
DB 508 GAACGTTGCGAGGTACAGAAACATCGTGAATTTGATGAGTGTTCATCCACTTT 567
QY 301 TCTCCAAATGTGCGCATCAGCTGTCTGTGCTTGAATCCTCGAGTCAAGTGTTCGGA 360
DB 568 TCTCCAAATGTGCGCATCAGCTGTCTGTGCTTGAATCCTCGAGTCAAGTGTTCGGA 627
QY 361 TTGCTCTTATTTCCAGTCAACGAGGTTGTTCCATGTGATGATACAGATTCGCGCA 420
DB 628 TTGCTCTTATTTCCAGTCAACGAGGTTGTTCCATGTGATGATACAGATTCGCGCA 687
QY 421 GATGTTTGGAGGCGCTGCTTCTTCAATGAGGGCTATGCTCATGCGGAGCTCAAA 480
DB 688 GATGTTTGGAGGCGCTGCTTCTTCAATGAGGGCTATGCTCATGCGGAGCTCAAA 747
QY 481 CCACGTATATTTGAGTGCAGAGAAATGATTTTAACTCATTTGACTT 540
DB 748 CCACGTATATTTGAGTGCAGAGAAATGATTTTAACTCATTTGACTT 807
QY 541 AGCTTCAAAAGAGGCAATCAGATGTAAATATTTCAAGACGGGTATCGGGCTCA 600
DB 808 AGCTTCAAAAGAGGCAATCAGATGTAAATATTTCAAGACGGGTATCGGGCTCA 867
QY 601 GAAGCAGAAATGCAAAATGCTTGGCCCAAGGCTGCGAGTATACAGATGTACC 660
DB 868 GAAGCAGAAATGCAAAATGCTTGGCCCAAGGCTGCGAGTATACAGATGTACC 927
QY 661 TCAGCTGTGATCTGTGAGAGCTTGAATCAATTTTCTGAAATGTTCTCAGGAATGAAA 720
DB 928 TCAGCTGTGATCTGTGAGAGCTTGAATCAATTTTCTGAAATGTTCTCAGGAATGAAA 987
QY 721 CTGAAACATACATCAGATCTCAGAAATGAGGCAAAACGTTCTGTATATTTGATCAC 780
DB 988 CTGAAACATACATCAGATCTCAGAAATGAGGCAAAACGTTCTGTATATTTGATCAC 1047
QY 781 ATATTGCGCAGTAAAGAGTGTGAATGCGCAATTCAGACTATACCTTAAAGACCTT 840
DB 1048 ATATTGCGCAGTAAAGAGTGTGAATGCGCAATTCAGACTATACCTTAAAGACCTT 1107
QY 841 ATCAAAAGATGCTTCAATGATTCAGAGCAAGAAATTCCTGCTGAATGCGATTGTGC 900
DB 1108 ATCAAAAGATGCTTCAATGATTCAGAGCAAGAAATTCCTGCTGAATGCGATTGTGC 1167
QY 901 AGCCCATTTCTTATGCAATCTTTTGGCCCTCATATTTGAAGATCTGGTCACTGCCACT 960
DB 1168 AGCCCATTTCTTATGCAATCTTTTGGCCCTCATATTTGAAGATCTGGTCACTGCCACT 1227
QY 961 CCAGTCTTAAAGATGCTGATGTGTGATGATGATTTCTTGGAGATGAAGAGAAAT 1020
DB 1228 CCAGTCTTAAAGATGCTGATGTGTGATGATGATTTCTTGGAGATGAAGAGAAAT 1287

```

```

QY 1021 GAGATGTTGTAGAAATGTAAAGAGAGTGTCAAAAATATGACCAAGTGTATCTTA 1080
DB 1288 GAGATGTTGTAGAAATGTAAAGAGAGTGTCAAAAATATGACCAAGTGTATCTTA 1347
QY 1081 CTTGTTCCAAAGAAATCTTGGGAGAGACAAAGCTTTGTTGAATGACAAATGCTGT 1140
DB 1348 CTTGTTCCAAAGAAATCTTGGGAGAGACAAAGCTTTGTTGAATGACAAATGCTGT 1407
QY 1141 GATTCCAAAGCTGCGCAGAAATTTACTGACTGGAAGATGTTTGAAGAGTTGTTG 1200
DB 1408 GATTCCAAAGCTGCGCAGAAATTTACTGACTGGAAGATGTTTGAAGAGTTGTTG 1467
QY 1201 GCTAATTTCTACCCCTGAGTCTTACAAAGAGGAGATCTGTATCAAACTTGTCTTA 1260
DB 1468 GCTAATTTCTACCCCTGAGTCTTACAAAGAGGAGATCTGTATCAAACTTGTCTTA 1527

```

RESULT 7
 AX16543 1260 bp DNA linear PAT 22-JUN-2001
 LOCUS
 DEFINITION Sequence 34 from Patent WO0138503.
 ACCESSION AX16543
 VERSION AX16543.1 GI:14546888

KEYWORDS
 SOURCE Homo sapiens (human)

ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
 1 Ploeman, G.D., Whyte, D., Manning, G.S., Sudarshan, S.S., Martinez, R.,
 Flanagan, P., and Clardy, D.S.
 Novel human protein kinases and protein kinase-like enzymes
 Patent: WO 0138503-A 34 31-MAY-2001;

JOURNAL
 Sugen, Inc. (US)

FEATURES
 source 1..1260
 Location/Qualifiers
 1..1260
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 99.6%; Score 1255.2; DB 6; Length 1260;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1257; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 ATGGCGGATCCGGCTGCGCTGCGGCGCGAGCGCGGCTTCTGAGGCTTCGCGG 60
DB 1 ATGGCGGATCCGGCTGCGCTGCGGCGCGAGCGCGGCTTCTGAGGCTTCGCGG 60
QY 61 CGGCTGTGCGAGGTACAGAGCGCTGTGGGTAGCGGCTCTCCGCTCGGTATCGGGTT 120
DB 61 CGGCTGTGCGAGGTACAGAGCGCTGTGGGTAGCGGCTCTCCGCTCGGTATCGGGTT 120
QY 121 CGCTGTGCGGGAACCTGGGCTGCGCCCGCGCGGCTCAAGCAATTCCTGCGCAGGA 180
DB 121 CGCTGTGCGGGAACCTGGGCTGCGCCCGCGCGGCTCAAGCAATTCCTGCGCAGGA 180
QY 181 ACCACCGGCGCTGCGGCTCTGCGCGCGAGTATGTTTCCGAAAGAGAGGCGCGCTG 240
DB 181 ACCACCGGCGCTGCGGCTCTGCGCGCGAGTATGTTTCCGAAAGAGAGGCGCGCTG 240
QY 241 GAACGTTGCGAGGTACAGAAACATCGTGAATTTGATGAGTGTTCATCCACTTT 300
DB 241 GAACGTTGCGAGGTACAGAAACATCGTGAATTTGATGAGTGTTCATCCACTTT 300
QY 301 TCTCCAAATGTGCGCATCAGCTGTCTGTGCTTGAATCCTCGAGTCAAGTGTTCGGA 360
DB 301 TCTCCAAATGTGCGCATCAGCTGTCTGTGCTTGAATCCTCGAGTCAAGTGTTCGGA 360
QY 361 TTGCTCTTATTTCCAGTCAACGAGGTTGTTCCATGTGATGATACAGATTCGCGCA 420
DB 361 TTGCTCTTATTTCCAGTCAACGAGGTTGTTCCATGTGATGATACAGATTCGCGCA 420

```



```

QY 841 ATCAAAAGCATGCTTATGATGATCCAGAGAGAAGATTCCTGCTGAAAATGGCATTTGTC 900
DB 1009 ATCAAAAAGCATGCTTATGATGATCCAGAGAGAAGATTCCTGCTGAAAATGGCATTTGTC 1068
QY 901 AGCCCATTTCTTAAAGCATTTCTTTTGGCCCTCATATTGAAAGATTGGTCAATGCTTCCACT 960
DB 1069 AGCCCATTTCTTAAAGCATTTCTTTTGGCCCTCATATTGAAAGATTGGTCAATGCTTCCACT 1128
QY 961 CCAGTGTCTAAGAGCTGCTGATGATGCTGATGATGATTTATCTTTGGAAATGAGAGAAATAT 1020
DB 1129 CCAGTGTCTAAGAGCTGCTGATGATGCTGATGATGATTTATCTTTGGAAATGAGAGAAATAT 1188
QY 1021 GAAGATGTTGTAGAGATGTAAGAGAGAGATGTCAAAAATATGAGACCAAGTGTATCTCTA 1080
DB 1189 GAAGATGTTGTAGAGATGTAAGAGAGAGATGTCAAAAATATGAGACCAAGTGTATCTCTA 1248
QY 1081 CTTGTCTCAAGAGAAAATCTCTGCGACAGACAAAGTCTTTGTGTGATGATGATGATGATGATGAT 1140
DB 1249 CTTGTCTCAAGAGAAAATCTCTGCGACAGACAAAGTCTTTGTGTGATGATGATGATGATGATGAT 1308
QY 1141 GATTCCAAAGCTGCGAGAAAATTAATGACTGAGAGAGATGTTGATGAGAAAGTTTGTGTG 1200
DB 1309 GATTCCAAAGCTGCGAGAAAATTAATGACTGAGAGAGATGTTGATGAGAAAGTTTGTGTG 1368
QY 1201 GCTACATTTTACCCGCTGAGTGCCTTCAAGAGGAGATATCTGTATCAAACTTTGCTTTAA 1260
DB 1369 GCTACATTTTACCCGCTGAGTGCCTTCAAGAGGAGATATCTGTATCAAACTTTGCTTTAA 1428

RESULT 9
BC058732 2041 bp mRNA linear ROD 30-JUN-2004
LOCUS Mus musculus kinase interacting with leukemia-associated gene 1
DEFINITION (strain:hm1), mRNA (CDNA clone MGC:64170 IMAGE:6414877), complete cds.
ACCESSION BC058732
VERSION BC058732.1 GI:37194892
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 2041)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Scapleorn, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Prange, C., Raha, S.S., Loguigliano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, D.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skliska, U., Small, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

TITLE
2 (bases 1 to 2041)
Strausberg, R.
Direct Submission
Submitted (24-SEP-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
REMARK

```

```

COMMENT
Contact: MGC help desk
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
CDNA Library Preparation: M. Bento Soares, University of Iowa
DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant.
Web site: http://genome.uiowa.edu
Contact: bent-soares@uiowa.edu; tom-casavant@uiowa.edu
Bonald, M.F., Akabou, I., Bair, T., Bair, J., Crouch, K., Davis, A., Fishler, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K., Scheetz, T., Smith, C., Snir, B., Tack, D., Trout, K., Walters, J., Casavant, T., Soares, M.B.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: http://image.lnlnl.gov
Series: Plate: Row: Column: 0
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6754447.
Location/Qualifiers
1..2041
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="MGC:64170 IMAGE:6414877"
/tissue_type="Brain, pool of enriched mouse brain tissues from 13.5, 14.5, 16.5 and 17.5 dpc"
/clone_lib="NIH BMAP_FRO"
/lab_host="DH10B"
/note="vector: pYX-A5C"
1..2041
/gene="K1st"
/note="synonym: K1S"
/db_xref="LOCUSID:16589"
/db_xref="MGI:1341908"
154..1413
/gene="K1st"
/codon_start=1
/product="kinase interacting with leukemia-associated gene (strain:hm1)"
/protein_id="AAHS8732.1"
/db_xref="GI:37194892"
/db_xref="LOCUSID:16589"
/db_xref="MGI:1341908"
/translation="MAGSGCAGWAEPPFLFLEAFRLMVOVSLGSSASVYRYVCCGTPGSPGALVKOFLPGTTGAASAAEYGRFRERALLBOLGHRALVTLYGVFTTIFSPNVSRCILBLIDVSVSLILYSSHSGSMMMOICARVDVLAFLHHEGVYADLKPNITKMSAENRFRKIDIGLSFKESNODVXYIOTDGVAPAPAEIONCIAOGLSDTECTSAVDLMSLGLILEMTSGMKLKTIVRSQEKANSSATIDHIFASKAVVNAATPAYH LKDLIKSMHDDPGRIIPAEMLCSPPFISIPAPHIDLVMLPTVPLRLVNLDDYLL ENEDYEVNVEVEKEECORQGVVSLVLPKKNPGQGVFEVYANAGDSKAOKLTLGR MEDGFVATFVFPISAVRYGLYQLL"

ORIGIN
Query Match 88.6%; Score 1116; DB 10; Length 2041;
Best Local Similarity 92.9%; Pred. No. 1.5e-278;
Matches 1170; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 1 ATGGCGGAGATCGGCGTGGCGCGCGGAGCGCGCGGCTTTCTGAGAGCGCTTGGG 60
DB 154 ATGGCGGAGATCGGCGTGGCGCGCGGAGCGCGCGGCTTTCTGAGAGCGCTTGGG 213
QY 61 CGGCTGTGGCAGGTACAGAGCGGTGTGGGTAGCGGCTCTCGGCTCGGTGTATCGGTT 120
DB 214 CGGCTGTGGCAGGTACAGAGCGGTGTGGGTAGCGGCTCTCGGCTCGGTGTATCGGTT 273
QY 121 CGCTGTGGCGCAACCTTGCTGCGCCCGCGCGCGCTTCAAGCAATTTCTTGGCCGACAGA 180
DB 274 CGGTGTGGCGCAACCTTGCTGCGCCCGCGCGCGCTTCAAGCAATTTCTTGGCCGACAGA 333
QY 181 ACCACCGGGGCTGCGGCGCTGCGCGCGAGTATGTTCCGCAAGAGAGAGGCGGCGCTG 240

```

Db	334	ACCAACCGGGGCTGCGGCTCCGGCCCGGAGTATGGTTTCCGCAAAAGAGGGCGGCGCTG	393
Qy	241	GAACAGTTGCAAGGCTCACAAGAAACATCGTGAATTGGTATGAGAGTGTTCATCACTTT	300
Db	394	GAGCAGTTGCAAGGGTCACAGAGAACATCGGACCTTATATATGAGGTCTTTACCATACATTC	453
Qy	301	TCTCCAAATGAGCCATCAACGCTGTCTGTGTCTTGAACCTCCGATGTCAAGTGTTCGGAA	360
Db	454	TCCTCCAAATGAGCCATCAACGCTGTCTGTGTCTTGAACCTCCGATGTCAAGTGTTCGGAA	513
Qy	361	TTGCTCTTATATTCACGTCAACCAAGGTTGTTCATATGGAATGATACAGATTGCCCGCA	420
Db	514	TTGCTTTTATATTCACGTCAATCAGGAGCTGTCTCATGTGGATGATACAGACTGTGCCGA	573
Qy	421	GATGTTTGGAGGCGCTTGTCTTTCTTCATCATGAGGGCTATGTCCATGCGGACTCAAA	480
Db	574	GATGTTCTGAGGGCCCTCGCTTTTCTTCAACCATGAGGGCTACGTCCATGCAAGACTCAAA	633
Qy	481	CCAGGTACATATTTGTGAGTGCACAGAAATGAATGTTTTTAACTCATTTGACTTTGGACTT	540
Db	634	CCAGCAAAATATCTGTGAGTGCAGAGAAATGATGTTTTTAAAGCTTATTTGACTTTGGACTT	693
Qy	541	AGCTTCAAGAAGGCATTCAGGATTTAAAGTATTTACAGACAGCGGATATCGGCTCCA	600
Db	694	AGCTTCAAGAAGGCATTCAGGACCTTAAGTATTTACAGACAGCGGATATCGGCTCCT	753
Qy	601	GAAGCAGAAATTCAGAAATTCCTTGGCCAGCGCTGGCTTCAGAGTGAATACAGATGTACC	660
Db	754	GAAGCAGAGCTGCAGAACTGCTGTGGCCAGCGGCTTCAGAGTGAATACAGATGTACC	813
Qy	661	TCAGCTGTGATCTGTGAGAGCTTAGGAATCATTTTACTGGAATGTTCTCAGAAATGAAA	720
Db	814	TCAGCTGTGATCTGTGAGAGCTTCAGAAATCATTTTACTGGAATGTTCTCAGGAATGAAG	873
Qy	721	CTGAAAACATACAGTCAGATCTCAGGAATGGAAGGCAAAACGTTCTGCTATTATTTGATAC	780
Db	874	CTGAAAACATACAGTGAATCTCAGGAATGGAAGGCAAAACGTTCTGCTATTATTTGATCAT	933
Qy	781	ATATTTGCGAGTAAAGCAGTGGTGAATCCCGCAATTCAGGCTATCAACTAAGACACTT	840
Db	934	ATATTTGCGAGTAAAGCAGTGGTGAATCCCGCAATTCAGGCTATCAACTAAGACACTT	993
Qy	841	ATCAAAAGCATGCTTCATGATGATTCAAAGCAGAGAAATTCCTGCTGAATGGCAATTGGC	900
Db	994	ATCAAAAGCATGCTTCATGATGATCCCGGCAAGAAATCCCTGCTGAAGTGGCAATTGGC	1053
Qy	901	AGCCCATTCCTTGAAGATTCCTTTTGGCCCTCATATTTGAAGATCTGGTATGCTCCCACT	960
Db	1054	AGCCCATTCCTTGAAGATTCCTTTTGGCCCTCATATTTGAAGATCTGGTATGCTCCGACT	1113
Qy	961	CCAGTGTCAAGACTGCTGAATGTGCTGTGATGATGTTATCTTGTGGAAATGAAGAGGAATAT	1020
Db	1114	CCAAGTGTCAAGGCTCTCTCAATGTGTGTGATGATGTTATCTTGAAGAAATGAAGATGAATAT	1173
Qy	1021	GAAGATGTTGTAGAAGATGTAAAGAGAGATGTCAAAAATATGACCAAGTGGTATCTCTA	1080
Db	1174	GAAGATGTTGTAGAAGATGTAAAGAGAGATGTCAAAAATATGACCAAGTGGTATCTCTG	1233
Qy	1081	CTTGTTCCAAAGGAAAAATCTGTGCGAGAGACAAAGCTTGTGTAGTATGCAAAAGCTGTGT	1144
Db	1234	CTTGTTCCAAAGGAAAAATCTGTGCGAGAGACAAAGCTTGTGTAGTATGCGAAACGCTGT	1293
Qy	1141	GATTCCAAAGCTGCGCAGAAAAATTAATGACTGGAGAGATTTGTGATGAGGAATTTGTGTG	1200
Db	1294	GATTCCAAAGCTGCGCAGAAAGTTGTGACTGGAGAGATTTGTGATGAGGAATTTGTGTG	1355
Qy	1201	GCTACATTCTACCGGCTGAGTGCTCTACAAAGGGGATATCTGTATCAAACTTGTCTTAA	1260
Db	1354	GCTACATTCTACCGGCTGAGTGCTCTACAAAGGGGATATCTTATCAAACTTGTCTTAA	1413

MPCOMMENT
Sequence split into 4 fragments LOCUS AC129141 Accession AC129141
Fragment Name Begin End
AC129141_0 1 110000
AC129141_1 100001 210000
AC129141_2 200001 310000
AC129141_3 300001 363206
Continuation (2 of 4) of AC129141 from base 100001 (AC129141 Rattus norvegicus clone CH23)

Query Match 88.3%; Score 1112.8; DB 2; Length 110000;
Best Local Similarity 92.7%; Pred. No. 2,3e-277;
Matches 1168; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

Qy	1	ATGGGGGATCCGGTGGCGCTTGGGGCGGGAGCCGCCGCTTTCTGAGAGCCTTCGGG	60
Db	70316	ATGGCGGGGCTCTGTGGCGGGGGGCGCGAGCGCCGCTTCTTGAGGGCTTTCGGG	70257
Qy	61	CGGCTGTGGCAGGTACAGACCCTCTGGGTACCGGCTCTCGGCTCGGATTCGGGTT	120
Db	70256	CGGCTGTGGCAGGTACAGACCCTCTAGGCAACGGCTCTCGGCTCGGATTCGGGTTG	70197
Qy	121	CGCTGTCGGCAACCTTGCTCGCTCCGCCCGCCCTCAAGCATGTTCTTGGCCGACAGA	180
Db	70196	CGCTGTCGGCAACCTTGCTCGCTCCGCCCGCCCTCAAGCATGTTCTTGGCCGACAGA	70137
Qy	181	ACCACCGGGGCTGGCGCTCTGCGCCGCGATATGTTTTCCGAAGAAGGGGGCGCTG	240
Db	70136	ACCACCGGGGCTGGCGCTCTGCGCTCGAGATATGTTTTCCGAAGAAGGGGGCGCTG	70077
Qy	241	GAACAGTTGACAGGCTCACAGAAACATCGATCTTGTATGAGTGTATTCAATCACTTT	300
Db	70076	GAGCAGTTGACAGGCTCACAGAAACATCGATCTTGTATGAGTGTATTCAATCACTTC	70017
Qy	301	TCTCCAAATGTCACATCAACGCTGTCTGTTGCTTGAACCTCTGGATCGATGTTTCGAA	360
Db	70016	TCTCCAAATGTCACATCAACGCTGTCTGTTGCTTGAACCTCTGGATCGATGTTTCGAA	69957
Qy	361	TTGCTCTATATTCGAGTCAACGAGGTTGTTCATATGTATGATATTAAGATTTGGCCGA	420
Db	69956	TTGCTCTATATTCGAGTCAACGAGGTTGTTCATATGTATGATATTAAGATTTGGCCGA	69897
Qy	421	GATGTTTGGAGGCCCTTGTCTTTCATCATGAGGCTATGTCATGCGGACTCAAA	480
Db	69896	GATGTTTGGAGGCCCTTGTCTTTCATCATGAGGCTATGTCATGCGGACTCAAA	69837
Qy	481	CCACGTAACATATTTGAGATGTCAGAGATGAATGTTTAACTAATGACTTTGGACTT	540
Db	69836	CCACGTAACATATTTGAGATGTCAGAGATGAATGTTTAACTAATGACTTTGGACTT	69777
Qy	541	AGCTTCAAGAAGGCAATCAGAGTGAATGAATTAATTCACAGACGGGTATCCGGCTCCA	600
Db	69776	AGCTTCAAGAAGGCAATCAGAGTGAATGAATTAATTCACAGACGGGTATCCAGCTCT	69717
Qy	601	GAAAGCAATTCGAAATTTGCTTGGCCGAGCGTGGCTTCAGAGTATACAGATGTACC	660
Db	69716	GAAAGCAATTCGAGAAATTTGCTTGGCCGAGCGCGCTTCAGAGTATACAGAGTACC	69657
Qy	661	TCAGCTGTGATCTGTGAGACCTTAGGAATCATTTTAACTGGAATGTTCTCAGAAATGAAA	720
Db	69656	TCAGCTGTGATCTGTGAGACCTTAGGAATCATTTTAACTGGAATGTTCTCAGAAATGAAA	69597
Qy	721	CTGAACCATACAGTCAGATCTCAGGAATGGAAGGCAACAGTCTGCTATTAATGATCAC	780
Db	69596	CTGAACCATACAGTCAGATCTCAGAGTGAAGGCAACAGTCTCTCTATTAATGATCAT	69537
Qy	781	ATATTTCGAGTAAAGCAGTGGTGAATGCCGAATTCAGCTATCACTTAAGAGACTT	840
Db	69536	ATATTTCGAGTAAAGCAGTGGTGAATGCCGAATTCAGCTATCACTTAAGAGACTT	69477
Qy	841	ATCAAAAGCATGCTTCATGATGATCCAAAGCAGAAATTCGCTGAATTTGGATTGTGC	900
Db	69476	ATCAAAAGCATGCTTCATGATGATCCCAAGCAGAAAGATTCCTGCTGAAGATGGCTTGTGC	69417

Db 999 CCAAGTCAGACTCCTCAATGCTGATGATGATCTATCTTGAATAAGATGATAT 1058
 QY 1021 GAAGATGTTAGAGATGTAAGAGAGAGTCAAAAATATGAGACCGATGATCTCA 1080
 Db 1059 GAAGATGTTAGAGATGTAAGAGAGAGTCAAAAATATGAGACCGATGATCTCA 1118
 QY 1081 CTGTTCCAAAGAAAATCTGCGAGAGCAAGTCTTTGTTGATGATGCAAAATGCTGT 1140
 Db 1119 CTGTTCCAAAGAAAATCTGCGAGAGCAAGTCTTTGTTGATGATGCAAAATGCTGT 1178
 QY 1141 GATTCCAAAGCTCGCAGAAAATTAATGATGATGATGATGATGATGATGATGATG 1200
 Db 1179 GATTCCAAAGCTCGCAGAAAATTAATGATGATGATGATGATGATGATGATGATG 1238
 QY 1201 GTTACATTCACCGCTGAGTGTGCTACAGAGGAGATGATGATGATGATGATGATG 1260
 Db 1239 GTTACATTCACCGCTGAGTGTGCTACAGAGGAGATGATGATGATGATGATGATG 1298

RESULT 12
 RNU70372 3244 bp mRNA linear ROD 08-SEP-1999
 LOCUS Rattus norvegicus PAM COOH-terminal Interactor protein 2 mRNA,
 DEFINITION complete cds.
 ACCSSION U70372.2 GI:5821767
 VERSION 1
 KEYWORDS Rattus norvegicus (Norway rat)
 SOURCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 3244)
 Alam,M.R., Caldwell,B.D., Johnson,R.C., Darlington,D.N., Maine,R.E.
 and Bipper,B.A.
 Novel proteins that interact with the COOH-terminal cytosolic
 routing determinants of an integral membrane peptide-processing
 enzyme
 J. Biol. Chem. 271 (45), 28636-28640 (1996)

JOURNAL
 MEDLINE 97067094
 PUBMED 8910496
 2 (bases 1 to 3244)
 Caldwell,B.D., Darlington,D.N., Penzes,P., Johnson,R.C.,
 Bipper,B.A. and Maine,R.E.
 The Novel Kinase P-CIP2 Interacts with the Cytosolic Routing
 Determinants of the Peptide Processing Enzyme Peptidylglycine
 alpha-Amidating Monooxygenase
 J. Biol. Chem. 274 (1999) In press
 3 (bases 1 to 3244)
 Alam,M.R., Caldwell,B.D., Johnson,R.C., Darlington,D.N., Maine,R.E.
 and Bipper,B.A.
 Direct Submission
 Submitted (11-SEP-1996) Neuroscience, Johns Hopkins University
 School of Medicine, 725 N. Wolfe Street, WBSB 902A, Baltimore, MD
 21205, USA
 4 (bases 1 to 3244)
 Caldwell,B.D., Darlington,D.N., Penzes,P., Johnson,R.C.,
 Bipper,B.A. and Maine,R.E.
 Direct Submission
 Submitted (03-SEP-1999) Neuroscience, Johns Hopkins University
 School of Medicine, 725 N. Wolfe Street, WBSB 902A, Baltimore, MD
 21205, USA
 REMARK
 COMMENT Sequence update by submitter
 FEATURES On Sep 3, 1999 this sequence version replaced gi:1698778.
 source Location/Qualifiers
 1..3244
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 167..1421
 /function="serine protein kinase"
 /note="P-CIP2"
 /codon_start=1

/product="PAM COOH-terminal Interactor protein 2"
 /protein_id="AAC53031.2"
 /db_xref="GI:5821768"
 /translation="MAGSCAMGAPPRLEAFGLMVOQSLSGSSASVYRACCG
 TPGSPGALKQPLPPTGTGAASABYRGSRKALBDQGRNITLVITGVTHFSP
 NVPSCLLELDVSEELVYSHQGSMMVIQICARDVLEALFLVHGVVHDLK
 PNIIIMSLENCEFKLIDFGLSFEKQGVYVQTDGYRPAEALQNLQALQSPT
 CTSAVDLMELGIIILEMFSGMKLKHTVRSQEKASSAIIDI PASKAVNNAIPAYH
 LADLIXSMHDDPSRRIPAMALCSPPRSIPRAPIHIDLVMLPTVRLINLVDDYTL
 ENDEYEDVEYVEKRCQRYGYSILVPEKNPFGQVFVEYANAGDSAAQKLTLGR
 MFDGRVVAATFPPLSAVKKGYLYDTLL"

ORIGIN

Query Match 87.9%; Score 1108; DB 10; Length 3244;
 Best Local Similarity 92.5%; Pred. No. 26-276;
 Matches 1165; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 1 ATGGCGGGAATCCGGCTGCGCTGCGGCGCGAGCCCGCTGTTCTGAGGCGCTTCGGG 60
 Db 162 ATGGCGGGAATCCGGCTGCGCTGCGGCGCGAGCCCGCTGTTCTGAGGCGCTTCGGG 221
 QY 61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTACGGCTCTCCGCTCGGTGTATCGGGTT 120
 Db 222 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTACGGCTCTCCGCTCGGTGTATCGGGGTG 281
 QY 121 CGCTGTGGGGAACCTGGCTGCGCCCGCGGCGGCTCAAGCAGTTCTTGCAGCAGGA 180
 Db 282 CGCTGTGGGGAACCTGGCTGCGCCCGCGGCGGCTCAAGCAGTTCTTGCAGCAGGA 341
 QY 181 ACCACCGGGGCTGCGGCTCTGCGCGCGAGTATGTTTCCGAAAGAGAGCGCGCTG 240
 Db 342 ACCACCGGGGCTGCGGCTCTGCGCGCGAGTATGTTTCCGAAAGAGAGCGCGCTG 401
 QY 241 GAACAGTTGAGGGGTACAGAAACATCGTACTTGTATGAGTGTATCAATCCACTTT 300
 Db 402 GAGCAGTTGAGGGGTACAGAAACATCGTACTTGTATGAGTGTATCAATCCACTTT 461
 QY 301 TCTCCAAATGTCATCAGCTGCTGTTGCTTGAACCTCGTAGATGATGATGATGATG 360
 Db 462 TCTCCAAATGTCATCAGCTGCTGTTGCTTGAACCTCGTAGATGATGATGATGATG 521
 QY 361 TTGCTCTTATTTTCAAGTACACAGGGTGTTCATGATGATGATGATGATGATGATG 420
 Db 522 TTGCTCTTATTTTCAAGTACACAGGGTGTTCATGATGATGATGATGATGATGATG 581
 QY 421 GATGTTTGAAGCCCTTCTTTCTTCAATGAGGCTATGTCATGCGGACCTCAAA 480
 Db 582 GATGTTTGAAGCCCTTCTTTCTTCAATGAGGCTATGTCATGCGGACCTCAAG 641
 QY 481 CCAAGTAAATTTGATGAGTGCAGAGATGAATGTTTAACTGATGATGATGATGATG 540
 Db 642 CCAAGTAAATTTGATGAGTGCAGAGATGAATGTTTAACTGATGATGATGATGATG 701
 QY 541 ACCTTCAAGAGAGCAATCAGATGTAAGTATTTTCAAGAGAGAGGATGCGGCTCA 600
 Db 702 ACCTTCAAGAGAGCAATCAGATGTAAGTATTTTCAAGAGAGAGGATGCGGCTCA 761
 QY 601 GAAGAGAAATTCGAAATTTGCTTGGCCCAAGGCTGCTGAGATGATGATGATGATG 660
 Db 762 GAAGAGAAATTCGAAATTTGCTTGGCCCAAGGCTGCTGAGATGATGATGATGATG 821
 QY 661 TCAAGCTTTGATCTGAGGCTTGAATCAATTTTCAAGAAATTTTCAAGAAATTTAA 720
 Db 822 TCAAGCTTTGATCTGAGGCTTGAATCAATTTTCAAGAAATTTTCAAGAAATTTAA 881
 QY 721 CTGAAACATACAGTACATCTCAGAAATGAGAGCAACAGTTCTGTATTTATGATCA 780
 Db 882 CTGAAACATACAGTACATCTCAGAAATGAGAGCAACAGTTCTGTATTTATGATCA 941
 QY 781 ATATTGCGAGTAAGAGAGTGTGAATGCGCAATTCAGCCTATACCTTAAGAGACTT 840
 Db 942 ATATTGCGAGTAAGAGAGTGTGAATGCGCAATTCAGCCTATACCTTAAGAGACTT 1001


```

Db      791 ATATTGGCCAGTAAACAGTGGTGAATGCGCAATTCAGCCATCATCCTCAGAGACCTT 850
Qy      841 ATCAAAAAGCATGCTTATGATGATCCAGACAGAAATTCCTGCTGAATGCAATTCGTC 900
Db      851 ATCAAAAAGCATGCTTATGATGATCCAGAGAGAAATTCCTGCTGAATGCAATTCGTC 910
Qy      901 AGCCCATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Db      911 AGCCCATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 970
Qy      961 CCAGTGTCAAGCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Db      971 CCAGTGTCAAGCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1030
Qy      1021 GAAGATGTTGAGAGATGTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Db      1031 GAAGATGTTGAGAGATGTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1090
Qy      1081 CTGTGTTCAAGGAAATCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Db      1091 CTGTGTTCAAGGAAATCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1150
Qy      1141 GATTCCAAAGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
Db      1151 GATTCCAAAGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1210
Qy      1201 GCTACATCTTACCCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1244
Db      1211 GCTACATCTTACCCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1254

```

RESULT 14

```

LOCUS      BC026046      1367 bp      mRNA      linear      PRI 30-JUN-2004
DEFINITION Homo sapiens kinase interacting with leukemia-associated gene
              (src1hmin), mRNA (cDNA clone MGC:39391 IMAGE:4649261), complete
              cds.
ACCESSION  BC026046
VERSION    BC026046.1  GI:19684094
KEYWORDS   MGC.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

REFERENCE

1 (bases 1 to 1367)
Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, K.H., Schmeider, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Uedlin, T.B., Toshikiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, K.D., Mullaly, S.J., Bosak, S.A., McKernan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Wortley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hui, X., Gibbs, R.A.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, B.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butcherfield, Y.S., Krzywinski, M.I., Skalska, U., Smallus, D.E.,
Scherer, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL PUBLISHED 12479932

REFERENCE 2 (bases 1 to 1367)

AUTHORS Straussberg, R.

JOURNAL TITLE Submitted (20-MAR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer

REMARK

COMMENT

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Crow, Allison Cloutier, Ruth
Featherstone, Malachi Griffith, Odi Griffith, Ran Guin, Nancy Liao,
Kim MacDonald, Amara Mason, Mike R. Mayo, Josh Moran, Ryan Morin,
Teika Olson, Diana Palmisani, Anca Petrescu, Ana Lisa Prabu,
Parvaneh Saeedi, JR Santos, Angelique Schermer, Ursula Skalska,
Duane Smalton, Jeff Stott, Miranda Tsai, George Yang, Jacques
Schein, Asim Siddiqui, Rob Holt, Marco Marra.

FEATURES

source

1..1367
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:39391 IMAGE:4649261"
/tissue_type="Eye, retinoblastoma"
/clone_lib="NIH MGC_16"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
1..1367
/gene="KIS"
/db_xref="LOCUSID:127933"
159..1193
/gene="KIS"
/codon_start=1
/product="kinase interacting src1hmin"
/protein_id="AAH26046.1"
/db_xref="GI:19684095"
/db_xref="LOCUSID:127933"
/translation="MAGSGCAGGAPPRFLAFLGLMVOYRSGSSASYRVACCG
NPGSPGALAKOFLPGTGAAGAASAEYFRERAALEQLOGRNIVLYGVFTIFSP
NVPSCRLELDVSELLLYSSHOGCSMMYIOHCARDVLEAFLHREGVHADLK
PNILMSAENECFLIDRGLSPKRGNDVYIOTDGYAPAPAIQNCIAQAGOSDTE
CTSAVDMSLGIILLHPSGKMKRTATVSGEKANSATTHIPASKAVVNAIATYH
LRLTLKSLHDDPSRITPAENALCSPPFSIPAPHTEDLVMLPTVLRNLNVLDYLL
ENEREYELIC"

CDS

gene

ORIGIN

Query Match 84.7%; Score 1067.2; DB 9; Length 1367;
Best Local Similarity 92.7%; Pred. No. 6; Se-266;
Matches 1168; Conservative 0; Mismatches 3; Indels 89; Gaps 1;
Qy 1 ATGGCGGATCCGAGTGGCGCTGAGCGCGAGCGCGCGCTTCTGAGAGCGCTTCGCG 60
Db 159 ATGGCGGATCCGAGTGGCGCTGAGCGCGAGCGCGCGCGCTTCTGAGAGCGCTTCGCG 218
Qy 61 CGGCTGTGGCAGTACAGAGCGCTGTGGGTAGCGGCTCTCCGCGCTCGGTGTAGGGTT 120
Db 219 CGGCTGTGGCAGTACAGAGCGCTGTGGGTAGCGGCTCTCCGCGCTCGGTGTAGGGTT 278
Qy 121 CGGCTGTGGCAGTACAGCGCTGCGCCCGCGCGCGCTCAAGCAGTCTTGGCGCGCAGA 180
Db 279 CGGCTGTGGCAGTACAGCGCTGCGCCCGCGCGCGCTCAAGCAGTCTTGGCGCGCAGA 338
Qy 181 ACCACCGGCGCTGCGCGCTTGGCGCGCGAGATGATGTTCCGAAAGAGAGCGCGCGCTG 240
Db 339 ACCACCGGCGCTGCGCGCTTGGCGCGCGAGATGATGTTCCGAAAGAGAGCGCGCGCTG 398

```

QY 241 GAACAGTTGACGGGTACAGAAAATCAGTGTGATGTGAGTGTGTTTCAATCCACTTT 300
DB 399 GAACAGTTGACGGGTACAGAAAATCAGTGTGATGTGAGTGTGTTTCAATCCACTTT 458
QY 301 TCTCCAAATGTGCGCATCAGCTGTCTGTGTTGAACTCTCTGGATGTCAAGTGTTCGGA 360
DB 459 TCTCCAAATGTGCGCATCAGCTGTCTGTGTTGAACTCTCTGGATGTCAAGTGTTCGGA 518
QY 361 TTGCTCTTATTTATTCAGTACACGAGGTGTGTTGATGTGATGTGATGAGATGTGCGCGA 420
DB 519 TTGCTCTTATTTATTCAGTACACGAGGTGTGTTGATGTGATGTGATGAGATGTGCGCGA 578
QY 421 GATGTTTGGAGGCGCTTGTCTTTCTTCAATCATAGAGGCTATGTCCATGCGGACTCAAA 480
DB 579 GATGTTTGGAGGCGCTTGTCTTTCTTCAATCATAGAGGCTATGTCCATGCGGACTCAAA 638
QY 481 CCAAGTAAATATTTGTGAGTGTGAGAGAAATGAAATGTTTAACTCATTTGAGACTT 540
DB 639 CCAAGTAAATATTTGTGAGTGTGAGAGAAATGAAATGTTTAACTCATTTGAGACTT 698
QY 541 AGCTTCAAAAGAGGCAATCAGATGTAAATATTTCAAGACAGGCTATGCGGCTCCA 600
DB 699 AGCTTCAAAAGAGGCAATCAGATGTAAATATTTCAAGACAGGCTATGCGGCTCCA 758
QY 601 GAAGCAGAAATTTGCAAAATTTGCTGGCCAGGCTGCGCTGAGATATCAGATGTACC 660
DB 759 GAAGCAGAAATTTGCAAAATTTGCTGGCCAGGCTGCGCTGAGATATCAGATGTACC 818
QY 661 TCAGCTGTGATGTGTGAGGCTTGAAGATCATTTTACTGGAAATGTTTCTCAGAAATGAAA 720
DB 819 TCAGCTGTGATGTGTGAGGCTTGAAGATCATTTTACTGGAAATGTTTCTCAGAAATGAAA 878
QY 721 CTGAAACATACAGTCTCAGATCTCAGAAATGAGGCAAAAGTTCTGCTACTATGTATCAC 780
DB 879 CTGAAACATACAGTCTCAGATCTCAGAAATGAGGCAAAAGTTCTGCTACTATGTATCAC 938
QY 781 ATATTTGCCAGTAAACAGTGTGAAATGCCGCAATTCAGGCTATCAGCTTAAGAGACTT 840
DB 939 ATATTTGCCAGTAAACAGTGTGAAATGCCGCAATTCAGGCTATCAGCTTAAGAGACTT 998
QY 841 ATCAAAAGCATGCTTATGATGATCCAGAGAGAAATTCCTGCTGAATGTGATGTGTC 900
DB 999 ATCAAAAGCATGCTTATGATGATCCAGAGAGAAATTCCTGCTGAATGTGATGTGTC 1058
QY 901 AGCCCATCTTGAATGCTTGTGCCCCCATATTTGAAGATGTGTCATGCTTCCACT 960
DB 1059 AGCCCATCTTGAATGCTTGTGCCCCCATATTTGAAGATGTGTCATGCTTCCACT 1118
QY 961 CCAGTGTCTAAGACTGCTGATGTGCTGATGATGATTAATCTTGGAAATGAGAGAAAT 1020
DB 1119 CCAGTGTCTAAGACTGCTGATGTGCTGATGATGATTAATCTTGGAAATGAGAGAAAT 1178
QY 1021 GAAGATGTGTGAGAAATGTAAGAGAGAGTCAAAAATATGAGACAGTGTGATCTCTA 1080
DB 1179 GAAG----- 1182
QY 1081 CTGTGTTCAAGAGAAATCTGTGAGAGAGCAAGTCTTGTGATGAGAAATGTCTGAT 1140
DB 1183 -----GTCCTTGTGATGAGTAAGCAATGTCTGAT 1209
QY 1141 GATTCGAAAGCTGCGAGAAATTAATGACTGAGAGAGATGTTGATGGGAATGTTGTGTG 1200
DB 1210 GATTCGAAAGCTGCGAGAAATTAATGACTGAGAGAGATGTTGATGGGAATGTTGTGTG 1269
QY 1201 GCTACATTTTACCCGCTGAGTGTCTTCAAGAGGAGATATCTGTATCAAACTTGTCTTAA 1260
DB 1270 GCTACATTTTACCCGCTGAGTGTCTTCAAGAGGAGATATCTGTATCAAACTTGTCTTAA 1329

```

```

RESULT 15
AK058195
LOCUS AK058195 1135 bp mRNA linear PRI 12-SEP-2003

```

```

DEFINITION Homo sapiens cDNA FLJ25466 fig, clone TST09304, highly similar to
ACCESSION AK058195
VERSION AK058195.1 GI:16554273
KEYWORDS oligo capping, fig (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Suzuki, O., Sasaki, N., Aotsuka, S., Shoji, T., Ichihara, T.,
Shionaka, N., Matsumoto, K., Hirano, M., Sato, S., Nomura, R.,
Yoshikawa, Y., Matsumura, Y., Moriya, S., Chiba, E., Momiyama, H.,
Onogawa, S., Kaeriyama, S., Sato, N., Matsumura, H., Takahashi, E.,
Katoka, R., Kuga, N., Kuroda, A., Satoh, I., Kamata, K., Takami, S.,
Terashima, Y., Watanabe, M., Suzuki, Y., Hata, H., Nakagawa, K.,
Mizuno, S., Morinaga, M., Kawamura, M., Sugiyama, T., Irie, R.,
Otsuki, T., Sato, H., Nishikawa, T., Sugiyama, A., Kawakami, B.,
Nagai, K., Isogai, T. and Sugano, S.
NEBO human cDNA sequencing project
Unpublished
TITLE NEBO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1135)
AUTHORS Sugano, S. and Suzuki, Y.
DIRECT SUBMISSION
TITLE Submitted (24-Oct-2001) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure, Human Genome
Center/ Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail: fliedna@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286,
Fax: 81-3-5449-5416)
COMMENT NEBO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction and 5'-end one pass sequencing: Institute of Medical
Science, University of Tokyo, Laboratory of Genome Structure, Human
Genome Center; 3'-end one pass sequencing: RAB; clone selection for
full insert sequencing: RAB and Helix Research Institute.
FEATURES
source
1. 1135
/organism="Homo sapiens"
/mol_type="cDNA"
/db_xref="taxon:9606"
/clone="TST09304"
/issue_type="testis"
/clone_11b="TST"
/note="Cloning vector: pMEL8FLJ3"
ORIGIN
Query Match 78.6%; Score 989.8; DB 9; Length 1135;
Best Local Similarity 99.8%; Pred. No. 86-246; Indels 0; Gaps 0;
Matches 991; Conservative 0; Mismatches 2;
QY 268 GTGACTTGTATGAGATGTTTCAATCCACTTTCTCCAAATGTGAGTGTCTG 327
DB 123 GTGACTTGTATGAGATGTTTCAATCCACTTTCTCCAAATGTGAGTGTCTG 182
QY 328 TTGCTGAACTCTGAGATGTCAAGTGTTCGAAATGCTTTATATTCAGTACCAAGGT 387
DB 183 TTGCTGAACTCTGAGATGTCAAGTGTTCGAAATGCTTTATATTCAGTACCAAGGT 242
QY 388 TGTTCATGTGATGATATACAGATTTGCGGCCGAGATGTTTGGAGGCGCTTTCTT 447
DB 243 TGTTCATGTGATGATATACAGATTTGCGGCCGAGATGTTTGGAGGCGCTTTCTT 302
QY 448 CATCATGAGGCGTATGTCCATGCGAGCTCAAAACAGCTTAATATTTGAGTGTGAGAG 507
DB 303 CATCATGAGGCGTATGTCCATGCGAGCTCAAAACAGCTTAATATTTGAGTGTGAGAG 362
QY 508 AATGAATGTTTAACTCATTTGACTTTGACTTTAGCTTTCAAGAGAGCAATCAGATGTA 567
DB 363 AATGAATGTTTAACTCATTTGACTTTGACTTTAGCTTTCAAGAGAGCAATCAGATGTA 422
QY 568 AAGTATATTCACAGAGCGGTATGCGGCTCCAGAGAGCAATTCGAAATTTGCTTGCGC 627

```

```
Db 423 AAGTATATTCAACAGACGGGTATCGGGCTCCAGAAAGCAGAAATTGCAAAATTGCTTGCC 482
Qy 628 CAGGCTGGCTGTGAGAGTATACAGAAATGACCTCAGCTGTGATGCTGTGAGCCTAGGA 687
Db 483 CAGGCTGGCTGTGAGAGTATACAGAAATGACCTCAGCTGTGATGCTGTGAGCCTAGGA 542
Qy 688 ATCATTTTACTGGAAATGTTCTCAGAAATGAAACTGAAACATACAGTCAATCTCAGAA 747
Db 543 ATCATTTTACTGGAAATGTTCTCAGAAATGAAACTGAAACATACAGTCAATCTCAGAA 602
Qy 748 TGGAAAGCAAAACAGTTCTGCTATTTATGATCAGATATTTGCCAGTAAAGCAGTGTGAAT 807
Db 603 TGGAAAGCAAAACAGTTCTGCTATTTATGATCAGATATTTGCCAGTAAAGCAGTGTGAAT 662
Qy 808 GCCGCAATTCAGCCCTATCACCCTAAGAGCCTTATCAAAAGCATGTCTTATGATGATCCA 867
Db 663 GCCGCAATTCAGCCCTATCACCCTAAGAGCCTTATCAAAAGCATGTCTTATGATGATCCA 722
Qy 868 AGCAGAAAGATTCCTGCTGAAATGCGATGTGCGAGCCCATCTTTAGCATTCCTTTGCC 927
Db 723 AGCAGAAAGATTCCTGCTGAAATGCGATGTGCGAGCCCATCTTTAGCATTCCTTTGCC 782
Qy 928 CCTCATATTTGAAGATCTGCTCATCTCCACTCCAGTGTAAAGTCTGCTGAATGTGCTG 987
Db 783 CCTCATATTTGAAGATCTGCTCATCTCCACTCCAGTGTAAAGTCTGCTGAATGTGCTG 842
Qy 988 GATGATGATTTATCTTGGGAATGAAGAAATGAAGATGTTGTAGAAGATGTAAAGAG 1047
Db 843 GATGATGATTTATCTTGGGAATGAAGAAATGAAGATGTTGTAGAAGATGTAAAGAG 902
Qy 1048 GAGTGTCAAAAATATGACCAAGTGTATCTCTACTTGTCCAAAAGAAAATCCTGSCAGA 1107
Db 903 GAGTGTCAAAAATATGACCAAGTGTATCTCTACTTGTCCAAAAGAAAATCCTGSCAGA 962
Qy 1108 GGAACAAGTCTTGTGTAGTATGCAAAATGCTGTGATTCAAAAGCTGSCAGAAATTAATG 1167
Db 963 GGAACAAGTCTTGTGTAGTATGCAAAATGCTGTGATTCAAAAGCTGSCAGAAATTAATG 1022
Qy 1168 ACTGGAAGATGTTGATGGGAAGTTGTGTGCTACATTCACCCGCTGAGTGCCTAC 1227
Db 1023 ACTGGAAGATGTTGATGGGAAGTTGTGTGCTACATTCACCCGCTGAGTGCCTAC 1082
Qy 1228 AAGAGGGATATCTGTATCAAACTTGCTTTAA 1260
Db 1083 AAGAGGGATATCTGTATCAAACTTGCTTTAA 1115
```

Search completed: November 29, 2004, 14:49:16
Job time : 5517 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2004, 11:23:25 ; Search time 653.5 Seconds
(without alignments)
10121.301 Million cell updates/sec

Title: US-10-798-532-1

Perfect score: 1260
Sequence: 1 atgsgcgagctccgctgcgc.....tgatcaaacctgtcttaaa 1260

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: geneeqn1980s:*
2: geneeqn1990s:*
3: geneeqn2000s:*
4: geneeqn2001s:*
5: geneeqn2001bs:*
6: geneeqn2002as:*
7: geneeqn2002bs:*
8: geneeqn2003as:*
9: geneeqn2003bs:*
10: geneeqn2003cs:*
11: geneeqn2003ds:*
12: geneeqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1260	100.0	1260	AAZ51355	AAZ51355 Human wil
2	1258.4	99.9	1260	AAZ51356	AAZ51356 Human tra
3	1258.4	99.9	2008	ABQ9383	ABQ9383 Human cod
4	1256.8	99.7	1260	AAI66829	AAI66829 Human pro
5	1256.8	99.7	2632	AAI66828	AAI66828 Human pro
6	1256.8	99.7	4065	AAH72714	AAH72714 Human cer
7	1256.8	99.7	4065	ADL45238	ADL45238 Human ova
8	1255.2	99.6	1260	AAZ51356	AAZ51356 Human tra
9	1255.2	99.6	1260	AAZ51356	AAZ51356 Human tra
10	1175	93.3	3998	ADQ24401	ADQ24401 Human kin
11	1108	87.9	3244	ADB52857	ADB52857 Primary r
12	458.6	36.4	529	AAH70764	AAH70764 Human cer
13	445.8	35.4	722	AAA02535	AAA02535 Human col
14	426.6	33.9	470	AAH71966	AAH71966 Human cer
15	375.4	29.8	461	AAH69101	AAH69101 Human cer
16	272.2	21.6	411	ABQ99225	ABQ99225 Human ORF
17	203.8	16.2	231	AAA01259	AAA01259 Human col
18	179	14.2	419	ADT37340	ADT37340 Human col
19	179	14.2	419	ADT37340	ADT37340 Human col
20	177.8	14.1	546	ABQ53127	ABQ53127 Oligonuc1
21	177.8	14.1	546	ABQ53126	ABQ53126 Oligonuc1

ALIGNMENTS

RESULT 1	AAZ51355	AAZ51355 standard; DNA; 1260 BP.	ABQ44465	ABQ44465 Oligonuc1
ID	AAZ51355	AAZ51355 standard; DNA; 1260 BP.	ABQ44464	ABQ44464 Oligonuc1
XX	AAZ51355;		ABQ53129	ABQ53129 Oligonuc1
XX			ABQ53128	ABQ53128 Oligonuc1
XX			ABQ44463	ABQ44463 Oligonuc1
XX			ABQ44462	ABQ44462 Oligonuc1
XX			ABQ44461	ABQ44461 Oligonuc1
XX			ABQ44460	ABQ44460 Oligonuc1
XX			ABQ44459	ABQ44459 Oligonuc1
XX			ABQ44458	ABQ44458 Oligonuc1
XX			ABQ44457	ABQ44457 Oligonuc1
XX			ABQ44456	ABQ44456 Oligonuc1
XX			ABQ44455	ABQ44455 Oligonuc1
XX			ABQ44454	ABQ44454 Oligonuc1
XX			ABQ44453	ABQ44453 Oligonuc1
XX			ABQ44452	ABQ44452 Oligonuc1
XX			ABQ44451	ABQ44451 Oligonuc1
XX			ABQ44450	ABQ44450 Oligonuc1
XX			ABQ44449	ABQ44449 Oligonuc1
XX			ABQ44448	ABQ44448 Oligonuc1
XX			ABQ44447	ABQ44447 Oligonuc1
XX			ABQ44446	ABQ44446 Oligonuc1
XX			ABQ44445	ABQ44445 Oligonuc1
XX			ABQ44444	ABQ44444 Oligonuc1
XX			ABQ44443	ABQ44443 Oligonuc1
XX			ABQ44442	ABQ44442 Oligonuc1
XX			ABQ44441	ABQ44441 Oligonuc1
XX			ABQ44440	ABQ44440 Oligonuc1
XX			ABQ44439	ABQ44439 Oligonuc1
XX			ABQ44438	ABQ44438 Oligonuc1
XX			ABQ44437	ABQ44437 Oligonuc1
XX			ABQ44436	ABQ44436 Oligonuc1
XX			ABQ44435	ABQ44435 Oligonuc1
XX			ABQ44434	ABQ44434 Oligonuc1
XX			ABQ44433	ABQ44433 Oligonuc1
XX			ABQ44432	ABQ44432 Oligonuc1
XX			ABQ44431	ABQ44431 Oligonuc1
XX			ABQ44430	ABQ44430 Oligonuc1
XX			ABQ44429	ABQ44429 Oligonuc1
XX			ABQ44428	ABQ44428 Oligonuc1
XX			ABQ44427	ABQ44427 Oligonuc1
XX			ABQ44426	ABQ44426 Oligonuc1
XX			ABQ44425	ABQ44425 Oligonuc1
XX			ABQ44424	ABQ44424 Oligonuc1
XX			ABQ44423	ABQ44423 Oligonuc1
XX			ABQ44422	ABQ44422 Oligonuc1
XX			ABQ44421	ABQ44421 Oligonuc1
XX			ABQ44420	ABQ44420 Oligonuc1
XX			ABQ44419	ABQ44419 Oligonuc1
XX			ABQ44418	ABQ44418 Oligonuc1
XX			ABQ44417	ABQ44417 Oligonuc1
XX			ABQ44416	ABQ44416 Oligonuc1
XX			ABQ44415	ABQ44415 Oligonuc1
XX			ABQ44414	ABQ44414 Oligonuc1
XX			ABQ44413	ABQ44413 Oligonuc1
XX			ABQ44412	ABQ44412 Oligonuc1
XX			ABQ44411	ABQ44411 Oligonuc1
XX			ABQ44410	ABQ44410 Oligonuc1
XX			ABQ44409	ABQ44409 Oligonuc1
XX			ABQ44408	ABQ44408 Oligonuc1
XX			ABQ44407	ABQ44407 Oligonuc1
XX			ABQ44406	ABQ44406 Oligonuc1
XX			ABQ44405	ABQ44405 Oligonuc1
XX			ABQ44404	ABQ44404 Oligonuc1
XX			ABQ44403	ABQ44403 Oligonuc1
XX			ABQ44402	ABQ44402 Oligonuc1
XX			ABQ44401	ABQ44401 Oligonuc1
XX			ABQ44400	ABQ44400 Oligonuc1
XX			ABQ44399	ABQ44399 Oligonuc1
XX			ABQ44398	ABQ44398 Oligonuc1
XX			ABQ44397	ABQ44397 Oligonuc1
XX			ABQ44396	ABQ44396 Oligonuc1
XX			ABQ44395	ABQ44395 Oligonuc1
XX			ABQ44394	ABQ44394 Arabidops

ALIGNMENTS	
Abq44465 Oligonuc1	Abq44465 Oligonuc1
Abq44464 Oligonuc1	Abq44464 Oligonuc1
Abq53129 Oligonuc1	Abq53129 Oligonuc1
Abq53128 Oligonuc1	Abq53128 Oligonuc1
Abq44463 Oligonuc1	Abq44463 Oligonuc1
Abq44462 Oligonuc1	Abq44462 Oligonuc1
Abq44461 Oligonuc1	Abq44461 Oligonuc1
Abq44460 Oligonuc1	Abq44460 Oligonuc1
Abq44459 Oligonuc1	Abq44459 Oligonuc1
Abq44458 Oligonuc1	Abq44458 Oligonuc1
Abq44457 Oligonuc1	Abq44457 Oligonuc1
Abq44456 Oligonuc1	Abq44456 Oligonuc1
Abq44455 Oligonuc1	Abq44455 Oligonuc1
Abq44454 Oligonuc1	Abq44454 Oligonuc1
Abq44453 Oligonuc1	Abq44453 Oligonuc1
Abq44452 Oligonuc1	Abq44452 Oligonuc1
Abq44451 Oligonuc1	Abq44451 Oligonuc1
Abq44450 Oligonuc1	Abq44450 Oligonuc1
Abq44449 Oligonuc1	Abq44449 Oligonuc1
Abq44448 Oligonuc1	Abq44448 Oligonuc1
Abq44447 Oligonuc1	Abq44447 Oligonuc1
Abq44446 Oligonuc1	Abq44446 Oligonuc1
Abq44445 Oligonuc1	Abq44445 Oligonuc1
Abq44444 Oligonuc1	Abq44444 Oligonuc1
Abq44443 Oligonuc1	Abq44443 Oligonuc1
Abq44442 Oligonuc1	Abq44442 Oligonuc1
Abq44441 Oligonuc1	Abq44441 Oligonuc1
Abq44440 Oligonuc1	Abq44440 Oligonuc1
Abq44439 Oligonuc1	Abq44439 Oligonuc1
Abq44438 Oligonuc1	Abq44438 Oligonuc1
Abq44437 Oligonuc1	Abq44437 Oligonuc1
Abq44436 Oligonuc1	Abq44436 Oligonuc1
Abq44435 Oligonuc1	Abq44435 Oligonuc1
Abq44434 Oligonuc1	Abq44434 Oligonuc1
Abq44433 Oligonuc1	Abq44433 Oligonuc1
Abq44432 Oligonuc1	Abq44432 Oligonuc1
Abq44431 Oligonuc1	Abq44431 Oligonuc1
Abq44430 Oligonuc1	Abq44430 Oligonuc1
Abq44429 Oligonuc1	Abq44429 Oligonuc1
Abq44428 Oligonuc1	Abq44428 Oligonuc1
Abq44427 Oligonuc1	Abq44427 Oligonuc1
Abq44426 Oligonuc1	Abq44426 Oligonuc1
Abq44425 Oligonuc1	Abq44425 Oligonuc1
Abq44424 Oligonuc1	Abq44424 Oligonuc1
Abq44423 Oligonuc1	Abq44423 Oligonuc1
Abq44422 Oligonuc1	Abq44422 Oligonuc1
Abq44421 Oligonuc1	Abq44421 Oligonuc1
Abq44420 Oligonuc1	Abq44420 Oligonuc1
Abq44419 Oligonuc1	Abq44419 Oligonuc1
Abq44418 Oligonuc1	Abq44418 Oligonuc1
Abq44417 Oligonuc1	Abq44417 Oligonuc1
Abq44416 Oligonuc1	Abq44416 Oligonuc1
Abq44415 Oligonuc1	Abq44415 Oligonuc1
Abq44414 Oligonuc1	Abq44414 Oligonuc1
Abq44413 Oligonuc1	Abq44413 Oligonuc1
Abq44412 Oligonuc1	Abq44412 Oligonuc1
Abq44411 Oligonuc1	Abq44411 Oligonuc1
Abq44410 Oligonuc1	Abq44410 Oligonuc1
Abq44409 Oligonuc1	Abq44409 Oligonuc1
Abq44408 Oligonuc1	Abq44408 Oligonuc1
Abq44407 Oligonuc1	Abq44407 Oligonuc1
Abq44406 Oligonuc1	Abq44406 Oligonuc1
Abq44405 Oligonuc1	Abq44405 Oligonuc1
Abq44404 Oligonuc1	Abq44404 Oligonuc1
Abq44403 Oligonuc1	Abq44403 Oligonuc1
Abq44402 Oligonuc1	Abq44402 Oligonuc1
Abq44401 Oligonuc1	Abq44401 Oligonuc1
Abq44400 Oligonuc1	Abq44400 Oligonuc1
Abq44399 Oligonuc1	Abq44399 Oligonuc1
Abq44398 Oligonuc1	Abq44398 Oligonuc1
Abq44397 Oligonuc1	Abq44397 Oligonuc1
Abq44396 Oligonuc1	Abq44396 Oligonuc1
Abq44395 Oligonuc1	Abq44395 Oligonuc1
Abq44394 Oligonuc1	Abq44394 Oligonuc1
Abq44393 Oligonuc1	Abq44393 Oligonuc1
Abq44392 Oligonuc1	Abq44392 Oligonuc1
Abq44391 Oligonuc1	Abq44391 Oligonuc1
Abq44390 Oligonuc1	Abq44390 Oligonuc1
Abq44389 Oligonuc1	Abq44389 Oligonuc1
Abq44388 Oligonuc1	Abq44388 Oligonuc1
Abq44387 Oligonuc1	Abq44387 Oligonuc1
Abq44386 Oligonuc1	Abq44386 Oligonuc1
Abq44385 Oligonuc1	Abq44385 Oligonuc1
Abq44384 Oligonuc1	Abq44384 Oligonuc1
Abq44383 Oligonuc1	Abq44383 Oligonuc1
Abq44382 Oligonuc1	Abq44382 Oligonuc1
Abq44381 Oligonuc1	Abq44381 Oligonuc1
Abq44380 Oligonuc1	Abq44380 Oligonuc1
Abq44379 Oligonuc1	Abq44379 Oligonuc1
Abq44378 Oligonuc1	Abq44378 Oligonuc1
Abq44377 Oligonuc1	Abq44377 Oligonuc1
Abq44376 Oligonuc1	Abq44376 Oligonuc1
Abq44375 Oligonuc1	Abq44375 Oligonuc1
Abq44374 Oligonuc1	Abq44374 Oligonuc1
Abq44373 Oligonuc1	Abq44373 Oligonuc1
Abq44372 Oligonuc1	Abq44372 Oligonuc1
Abq44371 Oligonuc1	Abq44371 Oligonuc1
Abq44370 Oligonuc1	Abq44370 Oligonuc1
Abq44369 Oligonuc1	Abq44369 Oligonuc1
Abq44368 Oligonuc1	Abq44368 Oligonuc1
Abq44367 Oligonuc1	Abq44367 Oligonuc1
Abq44366 Oligonuc1	Abq44366 Oligonuc1
Abq44365 Oligonuc1	Abq44365 Oligonuc1
Abq44364 Oligonuc1	Abq44364 Oligonuc1
Abq44363 Oligonuc1	Abq44363 Oligonuc1
Abq44362 Oligonuc1	Abq44362 Oligonuc1
Abq44361 Oligonuc1	Abq44361 Oligonuc1
Abq44360 Oligonuc1	Abq44360 Oligonuc1
Abq44359 Oligonuc1	Abq44359 Oligonuc1
Abq44358 Oligonuc1	Abq44358 Oligonuc1
Abq44357 Oligonuc1	Abq44357 Oligonuc1
Abq44356 Oligonuc1	Abq44356 Oligonuc1
Abq44355 Oligonuc1	Abq44355 Oligonuc1
Abq44354 Oligonuc1	Abq44354 Oligonuc1
Abq44353 Oligonuc1	Abq44353 Oligonuc1
Abq44352 Oligonuc1	Abq44352 Oligonuc1
Abq44351 Oligonuc1	Abq44351 Oligonuc1
Abq44350 Oligonuc1	Abq44350 Oligonuc1
Abq44349 Oligonuc1	Abq44349 Oligonuc1
Abq44348 Oligonuc1	Abq44348 Oligonuc1
Abq44347 Oligonuc1	Abq44347 Oligonuc1
Abq44346 Oligonuc1	Abq44346 Oligonuc1
Abq44345 Oligonuc1	Abq44345 Oligonuc1
Abq44344 Oligonuc1	Abq44344 Oligonuc1
Abq44343 Oligonuc1	Abq44343 Oligonuc1
Abq44342 Oligonuc1	Abq44342 Oligonuc1
Abq44341 Oligonuc1	Abq44341 Oligonuc1
Abq44340 Oligonuc1	Abq44340 Oligonuc1
Abq44339 Oligonuc1	Abq44339 Oligonuc1
Abq44338 Oligonuc1	Abq44338 Oligonuc1
Abq44337 Oligonuc1	Abq44337 Oligonuc1
Abq44336 Oligonuc1	Abq44336 Oligonuc1
Abq44335 Oligonuc1	Abq44335 Oligonuc1
Abq44334 Oligonuc1	Abq44334 Oligonuc1
Abq44333 Oligonuc1	Abq44333 Oligonuc1
Abq44332 Oligonuc1	Abq44332 Oligonuc1
Abq44331 Oligonuc1	Abq44331 Oligonuc1
Abq44330 Oligonuc1	Abq44330 Oligonuc1
Abq44329 Oligonuc1	Abq44329 Oligonuc1
Abq44328 Oligonuc1	Abq44328 Oligonuc1
Abq44327 Oligonuc1	Abq44327 Oligonuc1
Abq44326 Oligonuc1	Abq44326 Oligonuc1
Abq44325 Oligonuc1	Abq44325 Oligonuc1
Abq44324 Oligonuc1	Abq44324 Oligonuc1
Abq44323 Oligonuc1	Abq44323 Oligonuc1
Abq44322 Oligonuc1	Abq44322 Oligonuc1
Abq44321 Oligonuc1	Abq44321 Oligonuc1
Abq44320 Oligonuc1	Abq44320 Oligonuc1
Abq44319 Oligonuc1	Abq44319 Oligonuc1
Abq44318 Oligonuc1	Abq44318 Oligonuc1
Abq44317 Oligonuc1	Abq44317 Oligonuc1
Abq44316 Oligonuc1	Abq44316 Oligonuc1
Abq44315 Oligonuc1	Abq44315 Oligonuc1
Abq44314 Oligonuc1	Abq44314 Oligonuc1
Abq44313 Oligonuc1	Abq44313 Oligonuc1
Abq44312 Oligonuc1	Abq44312 Oligonuc1
Abq44311 Oligonuc1	Abq44311 Oligonuc1
Abq44310 Oligonuc1	Abq44310 Oligonuc1
Abq44309 Oligonuc1	Abq44309 Oligonuc1
Abq44308 Oligonuc1	Abq44308 Oligonuc1
Abq44307 Oligonuc1	Abq44307 Oligonuc1
Abq44306 Oligonuc1	Abq44306 Oligonuc1
Abq44305 Oligonuc1	Abq44305 Oligonuc1
Abq44304 Oligonuc1	Abq44304 Oligonuc1
Abq44303 Oligonuc1	Abq44303 Oligonuc1
Abq44302 Oligonuc1	Abq44302 Oligonuc1
Abq44301 Oligonuc1	Abq44301 Oligonuc1
Abq44300 Oligonuc1	Abq44300 Oligonuc1
Abq44299 Oligonuc1	Abq44299 Oligonuc1
Abq44298 Oligonuc1	Abq44298 Oligonuc1
Abq44297 Oligonuc1	Abq44297 Oligonuc1
Abq44296 Oligonuc1	Abq44296 Oligonuc1
Abq44295 Oligonuc1	Abq44295 Oligonuc1
Abq44294 Oligonuc1	Abq44294 Oligonuc1
Abq44293 Oligonuc1	Abq44293 Oligonuc1
Abq44292 Oligonuc1	Abq44292 Oligonuc1
Abq44291 Oligonuc1	Abq44291 Oligonuc1
Abq44290 Oligonuc1	Abq44290 Oligonuc1
Abq44289 Oligonuc1	Abq44289 Oligonuc1
Abq44288 Oligonuc1	Abq44288 Oligonuc1
Abq44287 Oligonuc1	Abq44287 Oligonuc1
Abq44286 Oligonuc1	Abq44286 Oligonuc1
Abq44285 Oligonuc1	Abq44285 Oligonuc1
Abq44284 Oligonuc1	Abq44284 Oligonuc1
Abq44283 Oligonuc1	Abq44283 Oligonuc1
Abq44282 Oligonuc1	Abq44282 Oligonuc1
Abq44281 Oligonuc1	Abq44281 Oligonuc1
Abq44280 Oligonuc1	Abq44280 Oligonuc1
Abq44279 Oligonuc1	Abq44279 Oligonuc1
Abq44278 Oligonuc1	Abq44278 Oligonuc1
Abq44277 Oligonuc1	Abq44277 Oligonuc1
Abq44276 Oligonuc1	Abq44276 Oligonuc1
Abq44275 Oligonuc1	Abq44275 Oligonuc1
Abq44274 Oligonuc1	Abq44274 Oligonuc1
Abq44273 Oligonuc1	Abq44273 Oligonuc1
Abq44272 Oligonuc1	Abq44272 Oligonuc1
Abq44271 Oligonuc1	Abq44271 Oligonuc1
Abq44270 Oligonuc1	Abq44270 Oligonuc1
Abq44269 Oligonuc1	Abq44269 Oligonuc1
Abq44268 Oligonuc1	Abq44268 Oligonuc1
Abq44267 Oligonuc1	Abq44267 Oligonuc1
Abq44266 Oligonuc1	Abq44266 Oligonuc1
Abq44265 Oligonuc1	Abq44265 Oligonuc1
Abq44264 Oligonuc1	Abq44264 Oligonuc1
Abq44263 Oligonuc1	Abq44263 Oligonuc1
Abq44262 Oligonuc1	Abq44262 Oligonuc1
Abq44261 Oligonuc1	Abq44261 Oligonuc1
Abq44260 Oligonuc1	Abq44260 Oligonuc1
Abq44259 Oligonuc1	Abq44259 Oligonuc1
Abq44258 Oligonuc1	Abq44258 Oligonuc1
Abq44257 Oligonuc1	Abq44257 Oligonuc1
Abq44256 Oligonuc1	Abq44256 Oligonuc1
Abq44255 Oligonuc1	Abq44255 Oligonuc1
Abq44254 Oligonuc1	Abq44254 Oligonuc1
Abq44253 Oligonuc1	Abq44253 Oligonuc1
Abq44252 Oligonuc1	Abq44252 Oligonuc1
Abq44251 Oligonuc1	Abq44251 Oligonuc1
Abq44250 Oligonuc1	Abq44250 Oligonuc1
Abq44249 Oligonuc1	Abq44249 Oligonuc1
Abq44248 Oligonuc1	Abq44248 Oligonuc1
Abq44247 Oligonuc1	Abq44247 Oligonuc1
Abq44246 Oligonuc1	Abq44246 Oligonuc1
Abq44245 Oligonuc1	Abq44245 Oligonuc1
Abq44244 Oligonuc1	Abq44244 Oligonuc1
Abq44243 Oligonuc1	Abq44243 Oligonuc1
Abq44242 Oligonuc1	Abq44242 Oligonuc1
Abq44241 Oligonuc1	Abq44241 Oligonuc1
Abq44240 Oligonuc1	Abq44240 Oligonuc1
Abq44239 Oligonuc1	Abq44239 Oligonuc1
Abq44238 Oligonuc1	Abq44238 Oligonuc1
Abq44237 Oligonuc1	Abq44237 Oligonuc1
Abq44236 Oligonuc1	Abq44236 Oligonuc1
Abq44235 Oligon	

XX The present DNA sequence encodes a wild type human KIS (hKIS), isolated
 CC from a human B-cell library, using a yeast two hybrid screening system.
 CC hKIS is a serine/threonine kinase, that acts as an inhibitory kinase of
 CC cyclin-dependent kinase inhibitor (CKI), p27. hKIS is a dominant gene
 CC that controls cell proliferation and is localised predominantly in the
 CC nucleus. It binds to CKI p27 and inhibits its ability to arrest cells in
 CC G1 phase. It has 99% homology to rat serine/threonine protein kinase KIS.
 CC The hKIS sequences are used to modulate cell proliferation and treat cell
 CC proliferative and vascular diseases. The polynucleotide sequence may be
 CC used in gene therapy to treat vascular disorders such as restenosis or
 CC atherosclerosis
 CC

SQ Sequence 1260 BP; 311 A; 280 C; 335 G; 334 T; 0 U; 0 Other;

Query Match 100.0%; Score 1260; DB 3; Length 1260;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGGATCCGCGTGCCTGGGCGCGAGACCGCGGCTTTCTGAGGCGCTTCGGG 60
 DB 1 ATGGCGGATCCGCGTGCCTGGGCGCGAGACCGCGGCTTTCTGAGGCGCTTCGGG 60
 QY 61 CGGCTGTGGGAGTACAGAGCCGCTGGGAGCGGCTCTCCGCTCGGTGATCGGGTT 120
 DB 61 CGGCTGTGGGAGTACAGAGCCGCTGGGAGCGGCTCTCCGCTCGGTGATCGGGTT 120
 QY 121 CGCTGTGGGCGCAACCGTGGCTGCGCCCGCGGCGCTCAAGAGATCTTCCGCGAGGA 180
 DB 121 CGCTGTGGGCGCAACCGTGGCTGCGCCCGCGGCGCTCAAGAGATCTTCCGCGAGGA 180
 QY 181 ACCAGCGGGGCTGCGGCTCTGCGCGCGAGATGATTTCCGCAAGAGAGCGGCGCTG 240
 DB 181 ACCAGCGGGGCTGCGGCTCTGCGCGCGAGATGATTTCCGCAAGAGAGCGGCGCTG 240
 QY 241 GAACATTTGCAAGGTGACAGAAACATCGTGAATTTGATGAGATGTTTCAATCCACTTT 300
 DB 241 GAACATTTGCAAGGTGACAGAAACATCGTGAATTTGATGAGATGTTTCAATCCACTTT 300
 QY 301 TCTCAAAATGTGCGATCAAGCTGTCTGTTGCTTGAACCTCTGAGATCAAGTGTTCGGA 360
 DB 301 TCTCAAAATGTGCGATCAAGCTGTCTGTTGCTTGAACCTCTGAGATCAAGTGTTCGGA 360
 QY 361 TTGCTCTTATATTCAGATCAAGAGGTTGTTCCAGTGTGATGATACAGCATTCGCGCGA 420
 DB 361 TTGCTCTTATATTCAGATCAAGAGGTTGTTCCAGTGTGATGATACAGCATTCGCGCGA 420
 QY 421 GATGTTTGGAGGCCCTTGCTTTTCTTCAATGAGAGGCTATTCATGCGGACTTCAAA 480
 DB 421 GATGTTTGGAGGCCCTTGCTTTTCTTCAATGAGAGGCTATTCATGCGGACTTCAAA 480
 QY 481 CCACGTAAATATTTGAGAGTGAAGAAATGATTTTAACTCAATGACTTTGAGCTT 540
 DB 481 CCACGTAAATATTTGAGAGTGAAGAAATGATTTTAACTCAATGACTTTGAGCTT 540
 QY 541 AGCTTCAAAAGAGCAATCAGATGTAAATATATTCAGACAGAGGATTCGGGCTCCA 600
 DB 541 AGCTTCAAAAGAGCAATCAGATGTAAATATATTCAGACAGAGGATTCGGGCTCCA 600
 QY 601 GAAGCAGATTTGCAAAATGCTTGGCCGAGGCTGCGCTGCAAGTATACAGAAATGACC 660
 DB 601 GAAGCAGATTTGCAAAATGCTTGGCCGAGGCTGCGCTGCAAGTATACAGAAATGACC 660
 QY 661 TCAGCTGTTGATCTGAGAGCTTGAAGATCAATTTACTGGAATGTTTCAGGAATGAAA 720
 DB 661 TCAGCTGTTGATCTGAGAGCTTGAAGATCAATTTACTGGAATGTTTCAGGAATGAAA 720
 QY 721 CTGAAACATACAGTCAAGATCTCAGGAATGGAAGCAAGTTCTCTATATATGATCAC 780
 DB 721 CTGAAACATACAGTCAAGATCTCAGGAATGGAAGCAAGTTCTCTATATATGATCAC 780
 QY 781 AATATTTCCAGTAAACAGATGTGATTCGCGCAATTCAGGCTTAAAGAGCTT 840
 DB 781 AATATTTCCAGTAAACAGATGTGATTCGCGCAATTCAGGCTTAAAGAGCTT 840

DB 781 AATATTTCCAGTAAACAGATGTGATTCGCGCAATTCAGGCTTAAAGAGCTT 840
 QY 841 ATCAAAAGCATCTTCATGATGATCCAGCAGAGAATTCCTGTGAATGSCATTGTC 900
 DB 841 ATCAAAAGCATCTTCATGATGATCCAGCAGAGAATTCCTGTGAATGSCATTGTC 900
 QY 901 AGCCCATTTTGAATTCCTTTGCCCCCTCAATTTGAAGATCTGATGCTTCCACT 960
 DB 901 AGCCCATTTTGAATTCCTTTGCCCCCTCAATTTGAAGATCTGATGCTTCCACT 960
 QY 961 CCAGTCTTAAAGCTGTAATGTGCTGATGATGATTTCTTGGAAATGAAAGGAATAT 1020
 DB 961 CCAGTCTTAAAGCTGTAATGTGCTGATGATGATTTCTTGGAAATGAAAGGAATAT 1020
 QY 1021 GAAGATGTTGTAAGATGTAAGAGAGAGTCTCAAAAATATGACCACTGATCTCTA 1080
 DB 1021 GAAGATGTTGTAAGATGTAAGAGAGAGTCTCAAAAATATGACCACTGATCTCTA 1080
 QY 1081 CTTGTTCCAAAGAAATCTCTGCGAGAGGACAAGTCTTGTGAGATGCAAAATCTGGT 1140
 DB 1081 CTTGTTCCAAAGAAATCTCTGCGAGAGGACAAGTCTTGTGAGATGCAAAATCTGGT 1140
 QY 1141 GATTCCAAAGCTGCGCAGAAATTAATCTGAAGAGATTTGATGGAAGTTGTG 1200
 DB 1141 GATTCCAAAGCTGCGCAGAAATTAATCTGAAGAGATTTGATGGAAGTTGTG 1200
 QY 1201 GCTACATTTACCCGCTGAGTCTCTACAAGAGGAGATCTGTATCAAACTTGTCTTAA 1260
 DB 1201 GCTACATTTACCCGCTGAGTCTCTACAAGAGGAGATCTGTATCAAACTTGTCTTAA 1260

RESULT 2
 AAZ51356
 ID AAZ51356 standard; DNA; 1260 BP.

XX AAZ51356;

DT 06-JUN-2000 (first entry)

XX

Human transdominant mutant serine/threonine kinase KIS (hKIS) gene.

XX

KM KIS; human; hKIS; serine/threonine kinase; cell proliferation; G1 phase;

KW transdominant gene; mutant; cyclin-dependent kinase inhibitor; CKI; p27;

KW modulator; treatment; cell proliferative disease; vascular disorder;

XX gene therapy; atherosclerosis; restenosis; ds.

OS Homo sapiens.

XX

Key Location/Qualifiers

FT CDS

1..1260

FT

/*tag= a

FT

/product= "Human transdominant mutant KIS protein"

FT

/transl_except= (pos:556..561, Asn)

FT

misc_feature

XX

161

XX

/*tag= b

XX

/note= "A to G substitution resulting in a Lys to Arg

XX

change in the amino acid sequence"

DR

MO20001165-A1.

XX

02-MAR-2000.

XX

20-AUG-1999; 99WO-US018903.

XX

21-AUG-1998; 98US-0097710P.

XX

(NABE/) NABEL G J.

XX

(NABE/) NABEL E G.

XX

Nabel GJ, Nabel EG;

XX

WPI: 2000-237648/20.

DR

P-PSDB; AAT70306.

XX Novel serine/threonine kinase hKIS polynucleotides and polypeptides used
PT for inhibiting the cyclin kinase inhibitor p27, and so alter cell
PT proliferation.

XX Claim 7, Page 59; 70pp; English.

XX The present DNA sequence encodes a transdominant mutant human KIS (hKIS),
CC constructed by site directed mutagenesis. A single nucleotide
CC substitution (A to G) results in a lysine to arginine change in the
CC protein sequence. hKIS is a serine/threonine kinase, that acts as an
CC inhibitory kinase of cyclin-dependent kinase inhibitor (CKI), p27. hKIS
CC controls cell proliferation and is localised predominantly in the
CC nucleus. It binds to CKI p27 and inhibits its ability to arrest cells in
CC G1 phase. The hKIS sequences are used to modulate cell proliferation and
CC treat cell proliferative and vascular diseases. The polynucleotide
CC sequence may be used in gene therapy to treat vascular disorders such as
CC stenosis or atherosclerosis

XX Sequence 1260 BP; 310 A; 280 C; 336 G; 334 T; 0 U; 0 Other;

XX Query Match 99.9%; Score 1258.4; DB 3; Length 1260;

XX Best Local Similarity 99.9%; Pred. No. 0;

XX Matches 1259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGGGATCCGGCTGCGCGCGCGGAGCGCGCGCGCTTTCTGAGGCGCTTCGGG 60
DB 1 ATGGCGGGATCCGGCTGCGCGCGCGGAGCGCGCGCGCTTTCTGAGGCGCTTCGGG 60
QY 61 CGGCTGTGCGAGGTACAGAGCCGTCTGGGTAGCGGCTCCGCTCGGCTGTATCGGGTT 120
DB 61 CGGCTGTGCGAGGTACAGAGCCGTCTGGGTAGCGGCTCCGCTCGGCTGTATCGGGTT 120
QY 121 CGCTGTGCGGCAACCTGCTGCGCGCGCGCGCGCGCTCAAGAGTTCTTCCCGCAGGA 180
DB 121 CGCTGTGCGGCAACCTGCTGCGCGCGCGCGCGCGCTCAAGAGTTCTTCCCGCAGGA 180
QY 181 ACCACCGGGGCTGCGCGCGCGCGCGAGTATGTTTCCGAAAGAGAGGGCGGCGCTG 240
DB 181 ACCACCGGGGCTGCGCGCGCGCGCGAGTATGTTTCCGAAAGAGAGGGCGGCGCTG 240
QY 241 GAACAGTTGCGAGGTACAGAAACATCGTATGTTATGAGTGTTCACATCCACTTT 300
DB 241 GAACAGTTGCGAGGTACAGAAACATCGTATGTTATGAGTGTTCACATCCACTTT 300
QY 301 TCTCCAAATGTCGATCAGCTGTCTGTGCTTGAACCTCTGGATGTCAGTGTTTCGAA 360
DB 301 TCTCCAAATGTCGATCAGCTGTCTGTGCTTGAACCTCTGGATGTCAGTGTTTCGAA 360
QY 361 TTGCTCTTATTTCCAGTACACGAGGTTTTCATGATGATGATAGATTTGCGGCCGA 420
DB 361 TTGCTCTTATTTCCAGTACACGAGGTTTTCATGATGATGATAGATTTGCGGCCGA 420
QY 421 GATGTTTGGAGGCGCTGCTTTCTTCATCATGAGGCTATGTCGATCGGACCTCAAA 480
DB 421 GATGTTTGGAGGCGCTGCTTTCTTCATCATGAGGCTATGTCGATCGGACCTCAAA 480
QY 481 CCAAGTAAATGTCGAGTCAAGATGAAATGTTTAACTCAATTGACTTGA 540
DB 481 CCAAGTAAATGTCGAGTCAAGATGAAATGTTTAACTCAATTGACTTGA 540
QY 541 AGCTTAAAGAGGCAATCAAGATGTAATGTAATTAATCAAGAGGATTCGGGCTCCA 600
DB 541 AGCTTAAAGAGGCAATCAAGATGTAATGTAATTAATCAAGAGGATTCGGGCTCCA 600
QY 601 GAAGCAGAAATGCAAAATGCTTGCGCCAGGCTGCGCTGCAAGATGATACAGATATACC 660
DB 601 GAAGCAGAAATGCAAAATGCTTGCGCCAGGCTGCGCTGCAAGATGATACAGATATACC 660
QY 661 TCAGCTGTGATGTGAGAGCTTGAAGATCAATTTTACTGGAATGTTTCAAGAAATGAAA 720
DB 661 TCAGCTGTGATGTGAGAGCTTGAAGATCAATTTTACTGGAATGTTTCAAGAAATGAAA 720

QY 721 CTGAATACATAGTCAGATCTCAGGAATGAGCAAGCAAGCTTCTGATTTATTTGATC 780
DB 721 CTGAATACATAGTCAGATCTCAGGAATGAGCAAGCAAGCTTCTGATTTATTTGATC 780
QY 781 ATATTGCGAGTAAAGAGGATGTAATCCGCAATTCAGCCCTATCAGCTTAAGAGACTT 840
DB 781 ATATTGCGAGTAAAGAGGATGTAATCCGCAATTCAGCCCTATCAGCTTAAGAGACTT 840
QY 841 ATCAAAAGATGCTTCAATGATGATCCAGCAAGAAATTCCTGCTGGAATGAGCAATTCG 900
DB 841 ATCAAAAGATGCTTCAATGATGATCCAGCAAGAAATTCCTGCTGGAATGAGCAATTCG 900
QY 901 AGCCATCTTTAGGATTCCTTTTCCCTCATATTTGAAGATCTGATGCTTCCACT 960
DB 901 AGCCATCTTTAGGATTCCTTTTCCCTCATATTTGAAGATCTGATGCTTCCACT 960
QY 961 CCAGTCTTAAGATCTGTAATGCTGATGATGATTTATCTTGGGAAATGAAGAGATAT 1020
DB 961 CCAGTCTTAAGATCTGTAATGCTGATGATGATTTATCTTGGGAAATGAAGAGATAT 1020
QY 1021 GAAGATGTTTGAAGATGTAATAAGAGATCTCAAAATATGACCAAGTGTATCTCTA 1080
DB 1021 GAAGATGTTTGAAGATGTAATAAGAGATCTCAAAATATGACCAAGTGTATCTCTA 1080
QY 1081 CTTGTTCCAAAGAAATCTTGGCAGAGACAAAGCTTTGTTGATGATGCAAAATGCTGGT 1140
DB 1081 CTTGTTCCAAAGAAATCTTGGCAGAGACAAAGCTTTGTTGATGATGCAAAATGCTGGT 1140
QY 1141 GATTCAAAGCTGCGCAGAAATTAAGTGAAGATGTTTGAAGAGAGGAGTTTGTG 1200
DB 1141 GATTCAAAGCTGCGCAGAAATTAAGTGAAGATGTTTGAAGAGAGGAGTTTGTG 1200
QY 1201 GCTAATTTCTACCGCTGAGCTTCAAGAGGAGATCTGATCAAACTTGTCTTAA 1260
DB 1201 GCTAATTTCTACCGCTGAGCTTCAAGAGGAGATCTGATCAAACTTGTCTTAA 1260

RESULT 3

AB099383
ID AB099383 standard; cDNA; 2008 BP.

XX AB099383;

XX 25-FEB-2003 (first entry)

XX Human coding sequence SEQ ID 116.

XX Human; expressed sequence tag; EST; haematopoietic disorder;

XX central nervous system disease; viral infection;

XX immune deficiency; immune disorder; bacterial infection; allergic; cancer;

XX fungal infection; autoimmune disorder; coagulation disorder; neutrotropic;

XX antiallergic; antiinflammatory; immunosuppressive; neuroprotective;

XX cytosolic; haemostatic; virucide; antibacterial; fungicide;

XX immunostimulant; cerebroprotective; gene therapy; gene; ss.

OS Homo sapiens.

PN WO200259260-A2.

XX 01-AUG-2002.

XX 16-NOV-2001; 2001WO-US042950.

XX 17-NOV-2000; 2000US-00714936.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
PI Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX WPI; 2002-590824/63.
XX N-PSDB; ABP64797.

XX New isolated polynucleotide, useful in research, diagnostic or
PT therapeutic methods, e.g. preventing or treating disorders involving
XX aberrant protein expression or biological activity.
XX Claim 1; SEQ ID NO 116; 394bp; English.

XX The present invention relates to novel human coding sequences (AB099268-
CC AB099608) and proteins (ABP64682-ABP65022). The sequences are useful in
CC therapeutic, diagnostic and research methods. The polynucleotides may be
CC used in the field of molecular biology as hybridisation probes, primers
CC for PCR, for chromosome and gene mapping, for the recombinant production
CC of protein, or in generation of anti-sense DNA or RNA. The
CC polynucleotides are useful in diagnostics as expressed sequence tags
CC (ESTs) for identifying expressed genes or for physical mapping of the
CC human genome. The proteins may be used as molecular weight markers, or as
CC nutritional sources or supplements. The proteins may be used to maintain
CC and expand cell population in a totipotent or pluripotent state
CC useful for re-engineering damaged or diseased tissues, transplantation,
CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The
CC polynucleotides and proteins are useful for preventing, treating or
CC ameliorating disorders involving aberrant protein expression or
CC biological activity, e.g. haematopoietic disorders, central/peripheral
CC nervous system diseases, mechanical and traumatic disorders, non-healing
CC wounds, immune deficiencies and disorders, infectious diseases caused by
CC viral, bacterial or fungal infection, autoimmune disorders, allergic
CC reactions and conditions, coagulation disorders, or cancer. The
CC polynucleotide sequences of the invention were assembled from ESTs
CC isolated mainly by sequencing by hybridisation, and in some cases,
CC sequences obtained from one or more public databases. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPD at
CC ftp.wipo.int/pub/published_pcr_sequences

SO Sequence 2008 BP; 491 A; 451 C; 488 G; 578 T; 0 U; 0 Other;

Query Match 99.9%; Score 1258.4; DB 6; Length 2008;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGGAAATCCGGCTGGCGCTGGGGGCGGAGACCGCGCTTTCTGAGAGCGCTTGGGG 60
DB 79 ATGGCGGAAATCCGGCTGGCGCTGGGGGCGGAGACCGCGCTTTCTGAGAGCGCTTGGGG 138
QY 61 CGGCTGTGGCAGGTACAGAGCCCTCTGGGTAGCGGCTCCCTCGCTGGGTATCCGGATT 120
DB 139 CGGCTGTGGCAGGTACAGAGCCCTCTGGGTAGCGGCTCCCTCGCTGGGTATCCGGATT 198
QY 121 CGGCTGTGGCAGGTACAGAGCCCTCTGGGTAGCGGCTCCCTCGCTGGGTATCCGGATT 180
DB 199 CGGCTGTGGCAGGTACAGAGCCCTCTGGGTAGCGGCTCCCTCGCTGGGTATCCGGATT 258
QY 181 ACCACCGGGGCTGGCGCTCTGGCGCGAGTATGTTTCCGCAAGAGAGGGCGGCGCTG 240
DB 259 ACCACCGGGGCTGGCGCTCTGGCGCGAGTATGTTTCCGCAAGAGAGGGCGGCGCTG 318
QY 241 GAACAGTTCCAGGCTCAGAGAAACATCGTGACTTTGTATGAGTGTTTCAATCACTTT 300
DB 319 GAACAGTTCCAGGCTCAGAGAAACATCGTGACTTTGTATGAGTGTTTCAATCACTTT 378
QY 301 TCTCCAAATGTGCTCAGTACAGGCTGTCTGTGCTTGAACCTCTGGATGTCACTGTTTGGAA 360
DB 379 TCTCCAAATGTGCTCAGTACAGGCTGTCTGTGCTTGAACCTCTGGATGTCACTGTTTGGAA 438
QY 361 TTGCTCTTATTTCCAGTACAGAGGTTGTTCATGTGATGATGATAGAGATTGGCGCGGA 420
DB 439 TTGCTCTTATTTCCAGTACAGAGGTTGTTCATGTGATGATGATAGAGATTGGCGCGGA 498
QY 421 GATGTTTGGAGGCCCTTGTCTTTCTTCATCATGAGGCTATGCTCATGCGGACTCAAA 480
DB 499 GATGTTTGGAGGCCCTTGTCTTTCTTCATCATGAGGCTATGCTCATGCGGACTCAAA 558
QY 481 CCAAGTAAATATTTGAGAGTGCAGAGAAATGATTTTAACTCATTTGAGACTT 540

DB 559 CCACCTTAATATTTGTGGAGTGTCAGAGAAATGATGATTTTAACTCATTTGAGACTT 618
QY 541 ACCTTCAAGAGGCAATCAGAGTGTAAATATTTCAAGACAGAGGATTCGGCTCA 600
DB 619 ACCTTCAAGAGGCAATCAGAGTGTAAATATTTCAAGACAGAGGATTCGGCTCA 678
QY 601 GAAGCAGATTTGCAAAATTTGCTTGGCCAGGCTGCTGCAGAGTGTATACAGATGTAC 660
DB 679 GAAGCAGATTTGCAAAATTTGCTTGGCCAGGCTGCTGCAGAGTGTATACAGATGTAC 738
QY 661 TAGCTGTGATCTGTGGAGCTTACAGATCATTTTACGTGAATTTCTCAGGAATGAA 720
DB 739 TAGCTGTGATCTGTGGAGCTTACAGATCATTTTACGTGAATTTCTCAGGAATGAA 798
QY 721 CTGAACATACAGTACAGATCTCAGAGATGAGCAACAGTTCTGATTTATTTGATAC 780
DB 799 CTGAACATACAGTACAGATCTCAGAGATGAGCAACAGTTCTGATTTATTTGATAC 858
QY 781 ATATTTGCCAGTAAAGCAGTGTGAAATGCCCAATTTCCAGCTTATCAGTAAAGACCTT 840
DB 859 ATATTTGCCAGTAAAGCAGTGTGAAATGCCCAATTTCCAGCTTATCAGTAAAGACCTT 918
QY 841 ATCAAAAGCATCTTCATGATGATTCAGACAGAGAAATTTCTGCTGAATGCAATTTGTC 900
DB 919 ATCAAAAGCATCTTCATGATGATTCAGACAGAGAAATTTCTGCTGAATGCAATTTGTC 978
QY 901 ACCCATTTCTTACGATTCCTTTGCCCTCATATTTGAAGATCTGATGCTTCCACT 960
DB 979 ACCCATTTCTTACGATTCCTTTGCCCTCATATTTGAAGATCTGATGCTTCCACT 1038
QY 961 CCAGTGTAAAGCTCTGAAATGCTGTGATGATGATTTATTTGGGAATGAGAGAAATAT 1020
DB 1039 CCAGTGTAAAGCTCTGAAATGCTGTGATGATGATTTATTTGGGAATGAGAGAAATAT 1098
QY 1021 GAAGATTTGTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 1099 GAAGATTTGTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1158
QY 1081 CTTGTTCCAAGAAATCTGTCGACAGAGCAAGTCTTTGTTGATGATGCAAAATGCTGAT 1140
DB 1159 CTTGTTCCAAGAAATCTGTCGACAGAGCAAGTCTTTGTTGATGATGCAAAATGCTGAT 1218
QY 1141 GATTCGAAGCTGCGCAGAAATTAATCTGAGAGAGATGTTGATGAGAAATGTTGTTG 1200
DB 1219 GATTCGAAGCTGCGCAGAAATTAATCTGAGAGAGATGTTGATGAGAAATGTTGTTG 1278
QY 1201 GCTAATTTCTACCGCTGATGCTTACAGAGGGAATATCTGATCAAACTTGTCTTAA 1260
DB 1279 GCTAATTTCTACCGCTGATGCTTACAGAGGGAATATCTGATCAAACTTGTCTTAA 1338

RESULT 4
AA166829
ID AA166829 standard; cDNA; 1260 BP.
XX
XX AA166829;
AC
XX
XX
DT 07-JAN-2002 (first entry)
XX
XX
DE Human protein kinase polypeptide 13887 coding sequence.
XX
XX Protein kinase; 3714; 16742; 23546; 13887; cancer; bone disorder; human;
XX cytosolic; antiinflammatory; immunosuppressive; cardiac; hepatocytic;
KM gene therapy; ss.
XX
XX Homo sapiens.
OS
XX
XX
FH Key Location/Qualifiers
FT 1..1260
FT CDS /tag= a
FT /product= "protein kinase 13887"
XX

PN WO200173050-A2.
 XX 04-OCT-2001.
 XX 23-MAR-2001; 2001WO-US009483.
 XX 24-MAR-2000; 2000US-0191846P.
 XX (MILL-) MILLENNIUM PHARM INC.
 XX Meyers R;
 XX WPI; 2001-611632/70.
 DR P-PSDB; AAG65767.
 XX
 PT New human protein kinase polypeptides, 3714, 16742, 23546 and 13887,
 PT useful in diagnosis of cancer or cellular proliferation or
 PT differentiation disorders and to screen for polypeptide modulators useful
 PT to treat such conditions.
 XX
 PS Claim 1; Fig 10A-C; 16pp; English.

CC The invention provides novel human protein kinase polypeptides, 3714,
 CC 16742, 23546 and 13887 and nucleic acid molecules encoding them. The
 CC protein kinase polypeptides can be expressed by standard recombinant
 CC methodology. 3714, 16742, 23546 or 13887 nucleic acids and polypeptides
 CC are useful for diagnostic and screening methods to identify subjects (at
 CC risk of) having cancer or cellular proliferation and/or differentiation
 CC disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and
 CC modulators are useful for the treatment of cancer, particularly colon
 CC cancer or cellular proliferation and/or differentiation disorders. Other
 CC disorders associated with 3714, 16742, 23546 or 13887 expression or
 CC activity that can be treated include bone related disorders, inflammatory
 CC diseases, autoimmune diseases, cardiovascular disorders and liver
 CC diseases. The present sequence represents a human protein kinase
 CC polypeptide 13887 coding sequence

Sequence 1260 BP; 312 A; 279 C; 334 G; 335 T; 0 U; 0 Other;

Query Match 99.7%; Score 1256.8; DB 4; Length 1260;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1258; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAGGCGGAGATCCGGCTGCGGCGGCGGCGGCGGCGGCTTCTGAGGCGCTTCGGG 60
 DB 1 AAGGCGGAGATCCGGCTGCGGCGGCGGCGGCGGCGGCGGCTTCTGAGGCGCTTCGGG 60
 QY 61 CGGCTGTGGCAGGTACAGAGCCGCTCTGGGTAGCGGCTCTCCGCTCGGTGTATCGGGTT 120
 DB 61 CGGCTGTGGCAGGTACAGAGCCGCTCTGGGTAGCGGCTCTCCGCTCGGTGTATCGGGTT 120
 QY 121 CGGTGTGGGCGCAACCTGTGCTGCGCCCGCGGCGGCTCAAGCATCTTCTGCGCGCAGGA 180
 DB 121 CGGTGTGGGCGCAACCTGTGCTGCGCCCGCGGCGGCTCAAGCATCTTCTGCGCGCAGGA 180
 QY 181 ACCACCGGGGCTGCGGCTCTGCGCGCGAGTATGTTTCCGCAAGAGAGGGGCGGCTG 240
 DB 181 ACCACCGGGGCTGCGGCTCTGCGCGCGAGTATGTTTCCGCAAGAGAGGGGCGGCTG 240
 QY 241 GAACAGTTGCAAGGTCACAGAAACATCTGACTTGTATGAGAGTCTTCAATCCACTTT 300
 DB 241 GAACAGTTGCAAGGTCACAGAAACATCTGACTTGTATGAGAGTCTTCAATCCACTTT 300
 QY 301 TCTCCAAATGTCACATCAGCTGTCTGTGTTGAACTCTCGATGTCAGTGTTCGAA 360
 DB 301 TCTCCAAATGTCACATCAGCTGTCTGTGTTGAACTCTCGATGTCAGTGTTCGAA 360
 QY 361 TTGCTCTTATATTCAGTACAGAGGTTGTTCCATGTGATGATAGATGAGGCGGCGGA 420
 DB 361 TTGCTCTTATATTCAGTACAGAGGTTGTTCCATGTGATGATAGATGAGGCGGCGGA 420
 QY 421 GATGTTTGGAGGCCCTTGCTTTCTTCAATCATGAGGCGTATGTCATGCGGACTCAAA 480
 DB 421 GATGTTTGGAGGCCCTTGCTTTCTTCAATCATGAGGCGTATGTCATGCGGACTCAAA 480

DB 421 GATGTTTGGAGGCCCTTGCTTTCTTCAATCATGAGGCGTATGTCATGCGGACTCAAA 480
 QY 481 CCAAGTAAATATTTGAGAGTGCAGAGAAATGATTTTAACTCATGACTTGGACTT 540
 DB 481 CCAAGTAAATATTTGAGAGTGCAGAGAAATGATTTTAACTCATGACTTGGACTT 540
 QY 541 AGCTTCAAGAAGGCAATCAGAGTGTAAAGTATATTCAGAGAGCGGCTATCGGCTCA 600
 DB 541 AGCTTCAAGAAGGCAATCAGAGTGTAAAGTATATTCAGAGAGCGGCTATCGGCTCA 600
 QY 601 GAAGCAGATTCGAAATTTGCTTGGCCCAAGGCTGCGCTGAGTGTATCAGATATGAC 660
 DB 601 GAAGCAGATTCGAAATTTGCTTGGCCCAAGGCTGCGCTGAGTGTATCAGATATGAC 660
 QY 661 TCAGCTGTGATCTGTGAGAGCTAGGAATCAATTTTCAAGTGAATGTCTCAGAAATGAA 720
 DB 661 TCAGCTGTGATCTGTGAGAGCTAGGAATCAATTTTCAAGTGAATGTCTCAGAAATGAA 720
 QY 721 CTGAAACATACAGTACAGTATCTCAGAGATGAAAGCAGATTCGCTATTAATGATAC 780
 DB 721 CTGAAACATACAGTACAGTATCTCAGAGATGAAAGCAGATTCGCTATTAATGATAC 780
 QY 781 ATATTTGCCAGTAAAGCAGTGTGATGAGCCGCAATTCAGCTTATCAGTAAAGACTT 840
 DB 781 ATATTTGCCAGTAAAGCAGTGTGATGAGCCGCAATTCAGCTTATCAGTAAAGACTT 840
 QY 841 ATCAAAGACATGCTTCAAGTATGTCAGAGAAATTCCTGCTGAAATGCAATGCTGTC 900
 DB 841 ATCAAAGACATGCTTCAAGTATGTCAGAGAAATTCCTGCTGAAATGCAATGCTGTC 900
 QY 901 AGCCATTTCTTGAAGATTCCTTTTCCCTCATATTTGAAGATCTGTATGCTTCCACT 960
 DB 901 AGCCATTTCTTGAAGATTCCTTTTCCCTCATATTTGAAGATCTGTATGCTTCCACT 960
 QY 961 CCAAGTCTAAGACTGCTGAATGCTGATGATGATATCTTGGAAATGAAGAGATAT 1020
 DB 961 CCAAGTCTAAGACTGCTGAATGCTGATGATGATATCTTGGAAATGAAGAGATAT 1020
 QY 1021 GAAGATGTTTGAAGATGTAAGAGAGTGTCAAAATATGAGCGAGTGTATCTCTA 1080
 DB 1021 GAAGATGTTTGAAGATGTAAGAGAGTGTCAAAATATGAGCGAGTGTATCTCTA 1080
 QY 1081 CTTGTTCCAAAGAAATCTGCGCAGAGACAAAGTCTTGTGTAGTATGCAAAATCTGCT 1140
 DB 1081 CTTGTTCCAAAGAAATCTGCGCAGAGACAAAGTCTTGTGTAGTATGCAAAATCTGCT 1140
 QY 1141 GATTCCAAAGCTGCGCAGAAATTAATCTGAGAGAGATGTTGATGAGAAAGTTGTG 1200
 DB 1141 GATTCCAAAGCTGCGCAGAAATTAATCTGAGAGAGATGTTGATGAGAAAGTTGTG 1200
 QY 1201 GGTACATTTCAACCGCTGAGTCCCTAAGAGGGATATCTGTATCAAACTTGTCTTAA 1260
 DB 1201 GGTACATTTCAACCGCTGAGTCCCTAAGAGGGATATCTGTATCAAACTTGTCTTAA 1260

RESULT 5
 ID AAI66828 strand; cDNA; 2622 BP.
 XX AAI66828;
 AC
 XX
 DT 07-JAN-2002 (first entry)
 XX
 XX Human protein kinase polypeptide 13887 encoding cDNA.
 DE
 XX Protein kinase; 3714; 16742; 23546; 13887; cancer; bone disorder; human;
 KW cytotoxic; antiinflammatory; immunosuppressive; cardiac; hepatotropic;
 KW gene therapy; ss.
 XX
 XX Homo sapiens.
 OS
 XX
 XX
 FH Key Location/Qualifiers
 FT 5'UTR 1..266

FT CDS /tag= a
FT 267.1529
FT /tag= b
FT /product= "protein kinase 13887"
FT /note= "coding sequence specifically claimed"
FT 3'UTR 1530.2622
FT /tag= c
XX WO200173050-A2.
XX 04-OCT-2001.
XX 23-MAR-2001; 2001WO-US009483.
XX 24-MAR-2000; 2000US-0191846P.
XX (MILL-) MILLENNIUM PHARM INC.
XX Meyers R;
XX WPI; 2001-611632/70.
XX P-PSDB; AAG65767.
XX
XX New human protein kinase polypeptides, 3714, 16742, 23546 and 13887,
PT useful in diagnosis of cancer or cellular proliferation or
PT differentiation disorders and to screen for polypeptide modulators useful
PT to treat such conditions.
XX
XX Claim 1; Fig 10A-C; 169pp; English.
XX
XX The invention provides novel human protein kinase polypeptides, 3714,
CC 16742, 23546 and 13887 and nucleic acid molecules encoding them. The
CC protein kinase polypeptides can be expressed by standard recombinant
CC methodology. 3714, 16742, 23546 or 13887 nucleic acids and polypeptides
CC are useful for diagnostic and screening methods to identify subjects (at
CC risk of) having cancer or cellular proliferation and/or differentiation
CC disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and
CC modulators are useful for the treatment of cancer, particularly colon
CC cancer or cellular proliferation and/or differentiation disorders. Other
CC disorders associated with 3714, 16742, 23546 or 13887 expression or
CC activity that can be treated include bone related disorders, inflammatory
CC disorders, autoimmune diseases, cardiovascular disorders and liver
CC diseases. The present sequence represents a human protein kinase
CC polypeptide 13887 encoding cDNA
XX
SQ Sequence 2622 BP; 652 A; 567 C; 627 G; 776 T; 0 U; 0 Other;
Query Match 99.7%; Score 1256.8; DB 4; Length 2622;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1258; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGCGGAGATCCGCGCTGCGCGCGCGCGAGACCCCGGCTTTTCTGAGGCGCTTCGGG 60
DB 267 ATGGCGGAGATCCGCGCTGCGCGCGCGCGAGACCCCGGCTTTTCTGAGGCGCTTCGGG 326
QY 61 CGGCTGTGGAGGTACAGAGCCGCTGTGGGTAGCGGCTCTCGGCTGTGTATCGGGTT 120
DB 327 CGGCTGTGGAGGTACAGAGCCGCTGTGGGTAGCGGCTCTCGGCTGTGTATCGGGTT 386
QY 121 CGGTGTGGGGAACCTGTGCTGCGCCCGCGGCGCCTCAAGAGTTCTTCCGCGAGGA 180
DB 387 CGGTGTGGGGAACCTGTGCTGCGCCCGCGGCGCCTCAAGAGTTCTTCCGCGAGGA 446
QY 181 ACCACCGGGGCTTGGGCTTCTGCGCGCGAGTATGTTTCCGAAAGAGAGGGCGGCTG 240
DB 447 ACCACCGGGGCTTGGGCTTCTGCGCGCGAGTATGTTTCCGAAAGAGAGGGCGGCTG 506
QY 241 GAACAGTTGAGGGGTACAGAAACATGTGACTTTGTATGAGTGTTTTCAATCACTTT 300
DB 507 GAACAGTTGAGGGGTACAGAAACATGTGACTTTGTATGAGTGTTTTCAATCACTTT 566
QY 301 TGTCCAAATGTGCAATCAGCTGTCTGTGCTTGAAGTCTCTGATATCAGTGTTCGGA 360
XX

DB 567 TCTCCAAATGTGCAATCAGCTGTCTGTGCTTGAAGTCTCTGATATCAGTGTTCGGA 626
QY 361 TTGCTCTTATATTCAGTACACAGAGGTTTTCATGTGATGATACAGATTCGCCGA 420
DB 627 TTGCTCTTATATTCAGTACACAGAGGTTTTCATGTGATGATACAGATTCGCCGA 686
QY 421 GATGTTTGGAGGCTTGTCTTTTCTTATCATCATGAGGGCTATGTCATCGGACTCAAA 480
DB 687 GATGTTTGGAGGCTTGTCTTTTCTTATCATCATGAGGGCTATGTCATCGGACTCAAA 746
QY 481 CCAAGTAAATATTTGTGAGTGCAGAGAAATGAAATGTTTAAATCATGACTTGTGACTT 540
DB 747 CCAAGTAAATATTTGTGAGTGCAGAGAAATGAAATGTTTAAATCATGACTTGTGACTT 806
QY 541 ACCTTCAAGAGGCAATCAGAGTAAAGTATATTCAGACAGAGGATTCGGCTCCA 600
DB 807 ACCTTCAAGAGGCAATCAGAGTAAAGTATATTCAGACAGAGGATTCGGCTCCA 866
QY 601 GAAGCAAGATTGCAAAATTTGCTTGGCCCAAGGCTGCGCTGAGATGATACAGATGACC 660
DB 867 GAAGCAAGATTGCAAAATTTGCTTGGCCCAAGGCTGCGCTGAGATGATACAGATGACC 926
QY 661 TGAGCTTTGATCTGTGAGGCTAGAGATCATTTTAACTGAGAAATGTTCTCAGAGAA 720
DB 927 TGAGCTTTGATCTGTGAGGCTAGAGATCATTTTAACTGAGAAATGTTCTCAGAGAA 986
QY 721 CTGAAACATACAGTACAGATCTCAGAGAAATGAGCAACAGTTCTGTATTTATGATAC 780
DB 987 CTGAAACATACAGTACAGATCTCAGAGAAATGAGCAACAGTTCTGTATTTATGATAC 1046
QY 781 ATATTTGCCAGTAAAGCAGTGTGATATGCGCAATTCAGCTTATCAGCTTAAAGACCTT 840
DB 1047 ATATTTGCCAGTAAAGCAGTGTGATATGCGCAATTCAGCTTATCAGCTTAAAGACCTT 1106
QY 841 ATCAAAAGATCTCTCAGATGATCCAGAGAGAAATTCGCTGGAATGAGCAATTTGTC 900
DB 1107 ATCAAAAGATCTCTCAGATGATCCAGAGAGAAATTCGCTGGAATGAGCAATTTGTC 1166
QY 901 AGCCCATCTTTAGCATTCCTTTTCCCTCATATTTGAAGATCTGTATGCTTCCACT 960
DB 1167 AGCCCATCTTTAGCATTCCTTTTCCCTCATATTTGAAGATCTGTATGCTTCCACT 1226
QY 961 CCAAGTCTAAGCTGCTGATATGCTGTGATGATATTTCTTTGGAAATGAGAGAAATAT 1020
DB 1227 CCAAGTCTAAGCTGCTGATATGCTGTGATGATATTTCTTTGGAAATGAGAGAAATAT 1286
QY 1021 GAAGATTTGTAGAGATGTAAGAGAGAGTCAAAATATGAGCAAGTGTATCTCTA 1080
DB 1287 GAAGATTTGTAGAGATGTAAGAGAGAGTCAAAATATGAGCAAGTGTATCTCTA 1346
QY 1081 CTTGTTCCAAAGAAATCTTGGCAGAGACAAATCTTTGTGATATGCAAAATGCTGCT 1140
DB 1347 CTTGTTCCAAAGAAATCTTGGCAGAGACAAATCTTTGTGATATGCAAAATGCTGCT 1406
QY 1141 GATTCCAAAGCTGCGCAGAAATTTACTGACTGGAAGATGTTTGAATGGAAGTTTGTG 1200
DB 1407 GATTCCAAAGCTGCGCAGAAATTTACTGACTGGAAGATGTTTGAATGGAAGTTTGTG 1466
QY 1201 GCTACATTTACCGGCTGAGTCTTCAAGAGGGGATTCGTATCAAACTTGTCTTAA 1260
DB 1467 GCTACATTTACCGGCTGAGTCTTCAAGAGGGGATTCGTATCAAACTTGTCTTAA 1526
RESULT 6
AAH72714
ID AAH72714 standard; cDNA; 4065 BP.
XX AAH72714;
AC
XX
DT 19-SBP-2001 (first entry)
XX
XX Human cervical cancer marker nucleic acid 3988.
XX

KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
 XX Homo sapiens.
 XX MO200142467-A2.
 XX
 PD 14-JUN-2001.
 XX
 PF 08-DEC-2000; 2000MO-US033312.
 XX
 PR 08-DEC-1999; 99US-0169681P.
 PR 21-DEC-1999; 99US-0171350P.
 PR 14-MAR-2000; 2000US-0189315P.
 PR 12-MAY-2000; 2000US-0203791P.
 PR 09-JUN-2000; 2000US-0210600P.
 PR 21-JUL-2000; 2000US-0220114P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PI Schlegel R, Deeds J, Berger A, Zhao X;
 XX WPI; 2001-375006/39.
 DR
 XX
 PT New isolated nucleic acid for diagnosing and treating cervical cancer and
 PT for assessing and detecting compounds for treating the cancer.
 XX
 PS Claim 1; Page 810-811; 1051pp; English.
 XX
 CC The invention relates to novel genes (AAH68727-AAH73383) associated with
 CC cervical cancer with cytostatic activity. The nucleic acids and encoded
 CC polypeptides are useful: to assess if a patient is afflicted with
 CC cervical cancer or has a pre-malignant condition; to monitor the
 CC progression of cervical cancer or a premalignant condition in a patient;
 CC and to select and/or assess the efficacy of a compound or therapy for
 CC inhibiting cervical cancer in a patient. The nucleic acids may also be
 CC useful for gene therapy
 CC
 SQ Sequence 4065 BP; 1068 A; 822 C; 905 G; 1265 T; 0 U; 5 Other;
 Query Match 99.7%; Score 1256.8; DB 4; Length 4065;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1258; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

DB 688 GATGTTTGGAGCCCTTCTTTCTTCATCATGAGGCTATGTCATGCGGACCTCAAA 747
 QY 481 CCAGCTAACATATATGTGAGAGTCAGAGATGATGATTTTAACTGATGACTTTGACTT 540
 DB 748 CACGTAACATATATGTGAGAGTCAGAGATGATGATTTTAACTGATGACTTTGACTT 807
 QY 541 AGCTTCAAAGAGGCAATCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 DB 808 AGCTTCAAAGAGGCAATCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 867
 QY 601 GAGCAGAAATTCGAAATTTGCTTGGCCAGAGCTGGCCCTGAGAGTATGATGATGATGAT 660
 DB 868 GAGCAGAAATTCGAAATTTGCTTGGCCAGAGCTGGCCCTGAGAGTATGATGATGATGAT 927
 QY 661 TAGCTGTTGATCTGAGAGCTGAGAAATCATTTTACTGAAATGTTCTCAGAAATGAAA 720
 DB 928 TAGCTGTTGATCTGAGAGCTGAGAAATCATTTTACTGAAATGTTCTCAGAAATGAAA 987
 QY 721 CTGAAACATACAGTACAGATCTCAGAAATGAGGCAAAAGTTCTGCTATTAATGATCAC 780
 DB 988 CTGAAACATACAGTACAGATCTCAGAAATGAGGCAAAAGTTCTGCTATTAATGATCAC 1047
 QY 781 ATATTTGCCAGTAAAGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 840
 DB 1048 ATATTTGCCAGTAAAGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1107
 QY 841 ATCAAAGATGCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
 DB 1108 ATCAAAGATGCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1167
 QY 901 ACCCATCTTTTGAAGATTCCTTTTCCCTCATATTTGAAGATCTGATGATGATGATGAT 960
 DB 1168 ACCCATCTTTTGAAGATTCCTTTTCCCTCATATTTGAAGATCTGATGATGATGATGAT 1227
 QY 961 CCAATGCTAAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
 DB 1228 CCAATGCTAAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1287
 QY 1021 GAAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
 DB 1288 GAAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1347
 QY 1081 CTTGTTCCAAAGAAATCTTGGCAGAGACAAAGTCTTTGTTGATGATGATGATGATGAT 1140
 DB 1348 CTTGTTCCAAAGAAATCTTGGCAGAGACAAAGTCTTTGTTGATGATGATGATGATGAT 1407
 QY 1141 GATTCGAAAGCTGCGCAGAAATTAAGTATGATGATGATGATGATGATGATGATGATGAT 1200
 DB 1408 GATTCGAAAGCTGCGCAGAAATTAAGTATGATGATGATGATGATGATGATGATGATGAT 1467
 QY 1201 GCTACATTTACCGCTGAGTCCCTACAAAGAGGATATCTGATCAAACTTGGCTTTAA 1260
 DB 1468 GCTACATTTACCGCTGAGTCCCTACAAAGAGGATATCTGATCAAACTTGGCTTTAA 1527
 RESULT 7
 ADL45238
 ID ADL45238 standard; DNA; 4065 BP.
 XX
 AC ADL45238;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 XX Human ovarian cancer DNA marker #19128.
 DE Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.
 XX
 OS Homo sapiens.
 XX
 FN MO200170979-A2.
 XX
 PD 27-SEP-2001.
 XX

PF 21-MAR-2001; 2001WO-US009126.
XX
XX 21-MAR-2000; 2000US-0191031P.
PR 25-MAY-2000; 2000US-0207124P.
PR 15-JUN-2000; 2000US-0211940P.
PR 07-JUL-2000; 2000US-0216820P.
PR 25-JUL-2000; 2000US-0220661P.
PR 21-DEC-2000; 2000US-0257672P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Lee J, Little J;
PI
DR WPI; 2001-611502/70.
XX
PT Novel isolated nucleic acid molecules (markers) overexpressed in ovarian
PT cancer cells as compared to their normal non-cancerous ovarian cells are
PT used to characterize stage, grade, histological type of ovarian cancer.
PS Disclosure; SEQ ID NO 19128; 106bp; English.
XX
XX The invention relates to nucleic acid markers which are overexpressed in
CC ovarian cancer cells as compared to their expression in normal (i.e. non-
CC cancerous) ovarian cells. The invention also relates to polypeptides
CC encoded by the markers, antibodies that selectively bind to the
CC polypeptides, a method of inhibiting ovarian cancer in a patient at risk
CC of developing ovarian cancer involving inhibiting expression of a gene
CC corresponding to a marker of the invention and a method of treating a
CC patient afflicted with ovarian cancer comprising providing to cells of
CC the patient an antisense oligonucleotide complementary to a marker of the
CC invention. The markers are useful for assessing if a patient is afflicted
CC with ovarian cancer, which involves comparing the level of expression of
CC a marker in a patient sample and a normal level of expression of the
CC marker in a control non-ovarian cancer sample. A difference between the
CC expression levels indicates ovarian cancer. The level of expression of a
CC marker corresponds to a secreted protein or to a transcribed
CC polynucleotide or its portion. The level of expression of the marker is
CC assessed by detecting the presence in the sample, a protein or protein
CC fragment corresponding to the marker. The presence of protein or protein
CC fragment is detected using an antibody that specifically binds with the
CC protein or protein fragment. Alternatively, the level of expression of
CC the marker is assessed by detecting the presence of a transcribed
CC polynucleotide which anneals with the marker or anneals with a portion of
CC the polynucleotide comprising the marker, under stringent conditions. The
CC marker is also used for monitoring the progression of ovarian cancer in a
CC patient which involves detecting expression of the marker in a patient
CC sample at a first point in time, repeating the method at a subsequent
CC time and comparing the level of expression. The method is carried out
CC using an ovarian tissue sample. A composition comprising a marker,
CC polypeptide or antibody of the invention is used to treat ovarian cancer.
CC This sequence represents a human ovarian cancer DNA marker of the
CC invention.
XX
XX Sequence 4065 BP; 1068 A; 822 C; 905 G; 1265 T; 0 U; 5 Other;
SQ
Query Match 99.7%; Score 1256.8; DB 5; Length 4065;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1258; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGCGGGAATCCGCTGCGCTTGGGGCGCGAGCCGCGCTTTTTCGAGGCTTTCCGG 60
DB 268 ATGGCGGGAATCCGCTGCGCTTGGGGCGCGAGCCGCGCTTTTTCGAGGCTTTCCGG 327
QY 61 CGGCTGTGGAGGTACAGAGCCGCTGGGTAGGCGGCTCCGCGCTCGGATGTGGGTT 120
DB 328 CGGCTGTGGAGGTACAGAGCCGCTGGGTAGGCGGCTCCGCGCTCGGATGTGGGTT 387
QY 121 CGCTGTGCGGCAACCTGTGCTGCCCGCGCGCCCTCAAGCAATTTCTTCCGCCAGGA 180
DB 388 CGCTGTGCGGCAACCTGTGCTGCCCGCGCGCCCTCAAGCAATTTCTTCCGCCAGGA 447
QY 181 ACCACGGGGGCTGCGGCTTCTGCCCGCGAGTATGTTTCCGCAAGAGAGGGCGGCTG 240
|||||

DB 448 ACCACGGGGGCTGCGGCTTCTGCCCGCGAGTATGTTTCCGCAAGAGAGGGCGGCTG 507
QY 241 GAACAGTTGCGAGGGTACAGAAACATCGTACTTTGTATGAGTGTTCATTCACCTT 300
DB 508 GAACAGTTGCGAGGGTACAGAAACATCGTACTTTGTATGAGTGTTCATTCACCTT 567
QY 301 TCTCCAAATGTGCATCAGCGTGTCTGTGCTTGAATCTCGGATGTCTAGTCTTCCGAA 360
DB 568 TCTCCAAATGTGCATCAGCGTGTCTGTGCTTGAATCTCGGATGTCTAGTCTTCCGAA 627
QY 361 TTGCTCTTATTTCCAGTACCAAGGTTTGTTCATGTGATGATATACAGATTCGCCCGCA 420
DB 628 TTGCTCTTATTTCCAGTACCAAGGTTTGTTCATGTGATGATATACAGATTCGCCCGCA 687
QY 421 GATGTTTGGAGGCCCTTCTTCTTCATCATGAGGGCTATGTCCATCGGACCTCAAA 480
DB 688 GATGTTTGGAGGCCCTTCTTCTTCATCATGAGGGCTATGTCCATCGGACCTCAAA 747
QY 481 CCACGTACATATTTGAGTGCAGAGAAATGATGTTTAACTCATTTGACTT 540
DB 748 CCACGTACATATTTGAGTGCAGAGAAATGATGTTTAACTCATTTGACTT 807
QY 541 AGCTTCAAAAGGCAATCAGATGTAAAGTATTCAGACGAGGGTATTCGGCTCCA 600
DB 808 AGCTTCAAAAGGCAATCAGATGTAAAGTATTCAGACGAGGGTATTCGGCTCCA 867
QY 601 GAAGCAGAAATTCGAAATGCTTGGCCGAGGCTGCGCTCAGAGTATCAGATGTAC 660
DB 868 GAAGCAGAAATTCGAAATGCTTGGCCGAGGCTGCGCTCAGAGTATCAGATGTAC 927
QY 661 TCAGCTGTGATCTGTGAGCCTTAGGAATCATTTTACTGAAATGTTCTCAGAAATGAA 720
DB 928 TCAGCTGTGATCTGTGAGCCTTAGGAATCATTTTACTGAAATGTTCTCAGAAATGAA 987
QY 721 CTGAAACATACAGTCAAGATCTCAGAAATGGAAGGCAACGTTTCGATTTTATGATAC 780
DB 988 CTGAAACATACAGTCAAGATCTCAGAAATGGAAGGCAACGTTTCGATTTTATGATAC 1047
QY 781 ATATTTCGAGTAAGAGTGTGTAATGCGCAATTCAGCTTACCTTAAGACCTT 840
DB 1048 ATATTTCGAGTAAGAGTGTGTAATGCGCAATTCAGCTTACCTTAAGACCTT 1107
QY 841 ATCAAAAGCATCTTCATGATATCCAAAGCAGAAATTCCTGTGAAATGSCATTTGTC 900
DB 1108 ATCAAAAGCATCTTCATGATATCCAAAGCAGAAATTCCTGTGAAATGSCATTTGTC 1167
QY 901 AGCCCATTTTTCATGATTCCTTTGCCCTCATTTTGAAGATCTGTATGCTTCCACT 960
DB 1168 AGCCCATTTTTCATGATTCCTTTGCCCTCATTTTGAAGATCTGTATGCTTCCACT 1227
QY 961 CCAGTGTAAAGCTGTGTAATGTGATGATGATTTCTTGGGAATGAAAGGAATAT 1020
DB 1228 CCAGTGTAAAGCTGTGTAATGTGATGATGATTTCTTGGGAATGAAAGGAATAT 1287
QY 1021 GAAGATGTTTGAAGATGTAAGAGAGAGTGTCAAAAATATAGCACTGTGTATCTTA 1080
DB 1288 GAAGATGTTTGAAGATGTAAGAGAGAGTGTCAAAAATATAGCACTGTGTATCTTA 1347
QY 1081 CTTGTTCCAAAGAAATTCCTGCGCAGAGACAAAGCTTGTGATGATGCAAAATCTGCT 1140
DB 1348 CTTGTTCCAAAGAAATTCCTGCGCAGAGACAAAGCTTGTGATGATGCAAAATCTGCT 1407
QY 1141 GATTCCAAAGCTGCGCAGAAATTAATCTGACTGGAAGATGTTTGAATGGAAGTTTGTG 1200
DB 1408 GATTCCAAAGCTGCGCAGAAATTAATCTGACTGGAAGATGTTTGAATGGAAGTTTGTG 1467
QY 1201 GCTTACATTTACCTGCTGAGTGCCTTACAAAGAGGGAATCTGTATCAAACTTGTCTTAA 1260
DB 1468 GCTTACATTTACCTGCTGAGTGCCTTACAAAGAGGGAATCTGTATCAAACTTGTCTTAA 1527
RESULT 8
AAS06734

AA06734 standard; cDNA; 1260 BP.
AA06734;
12-SEP-2001 (first entry)
Polynucleotide sequence encoding human protein kinase #34.
Human; protein kinase; PK; STK; cancer; cardiovascular disease;
metabolic disorder; immune related disease; neurological disorder;
neurodegenerative disorder; inflammatory disorder; infectious disease;
reproductive disorder; gene therapy; ss.
Homo sapiens.
WO200138503-A2.
31-MAY-2001.
22-NOV-2000; 2000WO-US032085.
24-NOV-1999; 99US-0167482P.
(SUGEN-) SUGEN INC.
PLOWMAN GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
Flanagan P, Clary D;
WPI; 2001-343950/36.
P-PSDB; AAU03534.
Nucleic acids encoding human kinase polypeptides, useful for preventing
diagnosing and/or treating e.g. cancer, immune, cardiovascular and
neural-associated diseases, and microbial infections.
Example 1; Fig 1; 433bp; English.
AA06701-AA06757 encode for novel human protein kinases #1-57. The novel
protein kinases have been identified as members of the tyrosine or
serine/threonine kinase (PTK and STK) families. The polynucleotides
encoding protein kinases and the polypeptides may be used in the
prevention, diagnosis and treatment of diseases associated with
inappropriate kinase expression. For example, they may be used to treat
cancers (especially cancers of hematopoietic origin), cardiovascular
disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
immune related diseases (e.g. rheumatoid arthritis), neurological
disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
disease (e.g. HIV) and reproductive disorders (e.g. infertility).
Additionally, polynucleotides encoding protein kinases may be used for
gene therapy and as DNA probes in diagnostic assays. The protein kinase
polypeptides may be used as antigens in the production of antibodies
against the protein kinases and in assays to identify modulators of
protein kinase expression and activity
Sequence 1260 BP; 311 A; 279 C; 335 G; 335 T; 0 U; 0 Other:
Query Match 99.6%; Score 1255.2; DB 4; Length 1260;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1257; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGGCGGAGATCCGGCTGGCGCTGGGGCGCGAGACCGCGGCTTTTCTGAGGCGCTTCGGG 60
DB 1 ATGGCGGAGATCCGGCTGGCGCTGGGGCGCGAGACCGCGGCTTTTCTGAGGCGCTTCGGG 60
QY CGGCTGTGCGAGATCAGAGACCGCTGTGGTAGGCGCTCTCGGCTCGAGTATCGAGATT 120
DB 61 CGGCTGTGCGAGATCAGAGACCGCTGTGGTAGGCGCTCTCGGCTCGAGTATCGAGATT 120
QY CGGCTGTGCGAGACCGCTGTGGCTGGCCCGCGGCGCTCAAGAGATTCTTTCGCGCAGGA 180
DB 121 CGGCTGTGCGAGACCGCTGTGGCTGGCCCGCGGCGCTCAAGAGATTCTTTCGCGCAGGA 180

QY 181 ACCACCGGGGCTGGGAGCTCTGCGCGCGAGTAGTGGTTTCCGCAAGAGAGGCGGCGCTG 240
DB 181 ACCACCGGGGCTGGGAGCTCTGCGCGCGAGTAGTGGTTTCCGCAAGAGAGGCGGCGCTG 240
QY 241 GAACAGTTGCGAGGCTCAGAGAAACATCGTGACTTTGTATGAGTGTTTACAATCCACTTT 300
DB 241 GAACAGTTGCGAGGCTCAGAGAAACATCGTGACTTTGTATGAGTGTTTACAATCCACTTT 300
QY 301 TCTCCAAATGTCGATCAACGCTGTCTGTGTAACCTCTGGATGTCAAGTGTTCGGAA 360
DB 301 TCTCCAAATGTCGATCAACGCTGTCTGTGTAACCTCTGGATGTCAAGTGTTCGGAA 360
QY 361 TTGCTCTTATATTCAGTCAACAGGGGTGTTCCATGTGATGATCAAGCATTCGCCGGA 420
DB 361 TTGCTCTTATATTCAGTCAACAGGGGTGTTCCATGTGATGATCAAGCATTCGCCGGA 420
QY 421 GATGTTTGGAGGCGCTGCTTTCTTCATCATGAGGGCTATGTCATGCGGACCTCAAA 480
DB 421 GATGTTTGGAGGCGCTGCTTTCTTCATCATGAGGGCTATGTCATGCGGACCTCAAA 480
QY 481 CCACGTAACATTTGTGAGTGCAGAGAAATGAATGTTTAACTCATTTGACTT 540
DB 481 CCACGTAACATTTGTGAGTGCAGAGAAATGAATGTTTAACTCATTTGACTT 540
QY 541 AGCTTCAAAAGGCAATCAGATGTAAAGTATATTCAGACAGAGGGATGCGGCTCCA 600
DB 541 AGCTTCAAAAGGCAATCAGATGTAAAGTATATTCAGACAGAGGGATGCGGCTCCA 600
QY 601 GAAGCAGATTGCAAAATTTGCTGGCCAGGCTGCGCAGAGTATCAGAAATGTACC 660
DB 601 GAAGCAGATTGCAAAATTTGCTGGCCAGGCTGCGCAGAGTATCAGAAATGTACC 660
QY 661 TAGCTGTGATCTGTGAGACCTTGAAGATCATTTTATCTGGAATGTTCTCAGAAATGAAA 720
DB 661 TAGCTGTGATCTGTGAGACCTTGAAGATCATTTTATCTGGAATGTTCTCAGAAATGAAA 720
QY 721 CTGAAACATACAGTCAAGATCTCAGGAATGGAAGGCAACAGTTCTGATATATGATCAC 780
DB 721 CTGAAACATACAGTCAAGATCTCAGGAATGGAAGGCAACAGTTCTGATATATGATCAC 780
QY 781 ATATTTGCGCAGTAAGACAGTGTGATGCGCAATTTCCAGCTTACCTTAAGAGACCTT 840
DB 781 ATATTTGCGCAGTAAGACAGTGTGATGCGCAATTTCCAGCTTACCTTAAGAGACCTT 840
QY 841 ATCAAAAGCATCTTCATGATATCAAGCAGAAATTTCTGTGAATGTCATTTGTGC 900
DB 841 ATCAAAAGCATCTTCATGATATCAAGCAGAAATTTCTGTGAATGTCATTTGTGC 900
QY 901 AGCCCATTTCTTGAATCTTTGCGGCTCATATTTGAAGATCTGTCATGCTCCACT 960
DB 901 AGCCCATTTCTTGAATCTTTGCGGCTCATATTTGAAGATCTGTCATGCTCCACT 960
QY 961 CCAAGTCTAAGACCTGTAATGCTGATGATGATATCTTGGAGATGAAAGAGAAATAT 1020
DB 961 CCAAGTCTAAGACCTGTAATGCTGATGATGATATCTTGGAGATGAAAGAGAAATAT 1020
QY 1021 GAAATGTTTGAAGATGTAAGAGAGAGTCAAAATATGACCAAGTGTATCTCTA 1080
DB 1021 GAAATGTTTGAAGATGTAAGAGAGAGTCAAAATATGACCAAGTGTATCTCTA 1080
QY 1081 CTGTGTTCAAAAGAAATCTGGCAGAGACAAGCTTTGTGAGATGCAATGCTGCT 1140
DB 1081 CTGTGTTCAAAAGAAATCTGGCAGAGACAAGCTTTGTGAGATGCAATGCTGCT 1140
QY 1141 GATTCAAAAGCTGCGAGAAATTAATGATGAGAGATGTTTGAATGGAAGTTTGTGTG 1200
DB 1141 GATTCAAAAGCTGCGAGAAATTAATGATGAGAGATGTTTGAATGGAAGTTTGTGTG 1200
QY 1201 GCTACATTTACCCGCTGAGTGCCTAACAAGAGGGAATCTGTATCAAACTTGTCTTAA 1260
DB 1201 GCTACATTTACCCGCTGAGTGCCTAACAAGAGGGAATCTGTATCAAACTTGTCTTAA 1260

```
RESULT 9
AADI8818
ID AADI8818 standard; cDNA; 1824 BP.
XX
AC AADI8818;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human kinase (PKIN) -3 cDNA.
XX
KW Human kinase; PKIN; gene therapy; adenocarcinoma; immune disorder; gout;
KW cancer; allergy; sarcoma; leukaemia; acquired immune deficiency syndrome;
KW AIDS; Addison's disease; microbial infection; inflammation; osteoporosis;
KW atherosclerosis; cardiovascular disease; myocardial infarction; anaemia;
KW myasthenia gravis; cirrhosis; cataract; growth and development disorder;
KW seizure disorder; pulmonary embolism; Gaucher's disease; lipid disorder;
KW lipid storage disease; Pick's disease; Tay-Sachs disease; renal disease;
KW asthma; obesity; restorative therapy; cytosolic; immunomodulatory;
KW antimicrobial; cardiovascular; antiinflammatory; vaccine; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS /tag= a
FT CDS /product= "Human PKIN-3 protein"
XX
PN MO2001B1555-A2.
XX
XX 01-NOV-2001.
XX
PF 20-APR-2001; 2001W0-US012992.
XX
XX 20-APR-2000; 2000US-0199021P.
XX
PR 28-APR-2000; 2000US-0200226P.
PR 05-MAY-2000; 2000US-0202339P.
PR 11-MAY-2000; 2000US-0203505P.
PR 18-MAY-2000; 2000US-0205564P.
PR 26-MAY-2000; 2000US-0207739P.
PR 01-JUN-2000; 2000US-0208795P.
XX
PI (INCY-) INCYTE GENOMICS INC.
PI
PI Yue H, Gandhi AR, Tribouley CM, Kearney L, Griffin JA, Nguyen DB,
PI Bandman O, Lu DM, Lal P, Burford N, Khan FA, Walla NK, Yao MG,
PI Patterson C, Burtill JD, Marcus GA, Zingler KA, Reardon SA, Lu Y,
PI Policky JL, Thornton M, Tang XT, Hafalia A, Elliott VS, Baughn MR,
PI Walsh RT, Ramkumar J, Borowsky ML, Au-Young J, Hillman JL,
PI Gururajan R;
XX
XX WPI; 2001-611740/70.
XX
XX P-PSDB; AAE11769.
XX
XX Human kinases and nucleic acids, useful for preventing diagnosing and
XX treating cancers, inflammation and immune disorders.
XX
XX Claim 5; Page 152; 166pp; English.
XX
XX The present invention relates to human kinases (PKIN) and the nucleic
XX acids encoding them. PKIN is used as vaccine and in gene therapy. PKIN is
XX used in the prevention, diagnosis and treatment of diseases cancers,
XX adenocarcinoma, leukaemia, sarcoma, immune disorder, Addison's disease,
XX acquired immune deficiency syndrome (AIDS), anaemia, asthma, allergies,
XX gout, microbial infections, cardiovascular disease and/or inflammation,
XX myasthenia gravis, atherosclerosis, cirrhosis, osteoporosis, myocardial
XX infection, cataract, growth and development disorder, seizure disorder,
XX pulmonary embolism, Gaucher's disease, lipid disorder, lipid storage
XX disease, Pick's disease, Tay-Sachs disease, renal disease and obesity.
XX PKIN may be used to treat disorders associated with decreased PKIN
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of PKIN by expressing inactive proteins or to
XX supplement the patient's own production of PKIN. PKIN nucleic acids may be
XX used to produce the PKIN polypeptide, by inserting the nucleic acids into
```

```
CC a host cell and culturing the cell to express the protein. PKIN nucleic
CC acid and its complementary sequences may also be used as DNA probes in
CC diagnostic assays to detect and quantitate the presence of similar
CC nucleic acid sequences in samples and therefore which patients may be in
CC need of restorative therapy. The present sequence is human PKIN-3 cDNA
XX
XX SQ Sequence 1824 BP; 434 A; 416 C; 472 G; 502 T; 0 U; 0 Other;
XX
XX Query Match 99.6%; Score 1255.2; DB 4; Length 1824;
XX Best Local Similarity 99.8%; Pred. No. 0;
XX Matches 1257; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
OY 1 ATGGCGGAGATCCGCTGCGCTGCGGAGCGGAGCGCGGCTTTCTGAGGCTTTGCGG 60
DB ATGGCGGAGATCCGCTGCGCTGCGGAGCGGAGCGCGGCTTTCTGAGGCTTTGCGG 228
OY CGGCTGTGGCAGTACAGAGCCGCTGTGGGTACGGCTCTCCGCTCGGTGTATGGGGTT 120
DB CGGCTGTGGCAGTACAGAGCCGCTGTGGGTACGGCTCTCCGCTCGGTGTATGGGGTT 288
OY CGCTGCTGGCGCAACCTTGCTGCCCCCGGCGCTTCAAGCAGTTCTTGCGCGCAGGA 180
DB CGCTGCTGGCGCAACCTTGCTGCCCCCGGCGCTTCAAGCAGTTCTTGCGCGCAGGA 348
OY ACCACCGGGGCTGCGGCTCTGCGCCGCAATATGTTTCCGAAAGAGGCGCGCTGT 240
DB ACCACCGGGGCTGCGGCTCTGCGCCGCAATATGTTTCCGAAAGAGGCGCGCTGT 408
OY GAACAGTTGCAAGGTACAGAAACATTCGACTTTGTATGAGTGTATCAATCACTTT 300
DB GAACAGTTGCAAGGTACAGAAACATTCGACTTTGTATGAGTGTATCAATCACTTT 468
OY TCTCCAAATGTCCATCAACGCTGTCTGTTGCTTGAACCTCGATGTCAGTGTTCGGA 360
DB TCTCCAAATGTCCATCAACGCTGTCTGTTGCTTGAACCTCGATGTCAGTGTTCGGA 528
OY TGTGCTTTATTTTCAAGTACAGAGGTTGTTTCATGTGATATACAGATGTGCCGA 420
DB TGTGCTTTATTTTCAAGTACAGAGGTTGTTTCATGTGATATACAGATGTGCCGA 588
OY GATGTTTGGAGCCCTGTTTCTTCATATGAGGGGTATGTCAGTCCGACCTCAAA 480
DB GATGTTTGGAGCCCTGTTTCTTCATATGAGGGGTATGTCAGTCCGACCTCAAA 648
OY CCACGTAACATTTTGTGAGTGCAGAGATGATGTTTAACTGATGACTTTGGACTT 540
DB CCACGTAACATTTTGTGAGTGCAGAGATGATGTTTAACTGATGACTTTGGACTT 708
OY AGCTTCAAAAGGCAATCAGATGTAAAGTATTTTCAACAGACGCGGTATCGGCTCCA 600
DB AGCTTCAAAAGGCAATCAGATGTAAAGTATTTTCAACAGACGCGGTATCGGCTCCA 768
OY GAAGCAGATTGCAAAATTTGTTGGCCGAGGCTGCGCAGAGTATCAGATATGTAAC 660
DB GAAGCAGATTGCAAAATTTGTTGGCCGAGGCTGCGCAGAGTATCAGATATGTAAC 828
OY TCAGCTGTTGATCTGTGAGCTTGAAGATCAATTTTACGGAATATTTTCAGGAATGAA 720
DB TCAGCTGTTGATCTGTGAGCTTGAAGATCAATTTTACGGAATATTTTCAGGAATGAA 888
OY CTGAAACATACATGATCTGAGATCTGAGAAATGGAAGCAAGTTCTGCTATTTATGATC 780
DB CTGAAACATACATGATCTGAGATCTGAGAAATGGAAGCAAGTTCTGCTATTTATGATC 948
OY AATTTTGCAGTAAAGCAGTGTGAATGCGGAATTTTCAAGCTTACCTTAAGAGACTT 840
DB AATTTTGCAGTAAAGCAGTGTGAATGCGGAATTTTCAAGCTTACCTTAAGAGACTT 1008
OY ATCAAAAGCATCTTCATGATGATCCAAAGCAGAAATTTCTGCTGAATGCAATTTGTC 900
DB ATCAAAAGCATCTTCATGATGATCCAAAGCAGAAATTTCTGCTGAATGCAATTTGTC 1068
OY ATCAAAAGCATCTTCATGATGATCCAAAGCAGAAATTTCTGCTGAATGCAATTTGTC 960
DB ATCAAAAGCATCTTCATGATGATCCAAAGCAGAAATTTCTGCTGAATGCAATTTGTC 960
```



```
|||||
Db 1069 AGCCATCTTATAGATCTTTGGCCCTCATATGAAGATCTGATGCTCCCACT 1128
Qy 961 CCAGTCTAAGACTGCTGATATGCTGATGATGATATCTTGGGAATGAAGGAATAT 1020
Db 1129 CCAGTCTAAGACTGCTGATATGCTGATGATGATATCTTGGGAATGAAGGAATAT 1188
Qy 1021 GAAGATGTTGTAAGATGTAAGAGAGATGTCAAAATATGAGACAGTGGATCTCTA 1080
Db 1189 GAGATGTTGTAAGATGTAAGAGAGATGTCAAAATATGAGACAGTGGATCTCTA 1248
Qy 1081 CTTGTTCCAAAGAAATCTGCGCAGAGACAAGTCTTTGTTGATGATGCAATGCTGCT 1140
Db 1249 CTTGTTCCAAAGAAATCTGCGCAGAGACAAGTCTTTGTTGATGATGCAATGCTGCT 1308
Qy 1141 GATTCGAAAGCTGCGGAGAAATTAATGATGATGAGAGATGTTGATGAGAAATTTGTTG 1200
Db 1309 GATTCGAAAGCTGCGGAGAAATTAATGATGATGAGAGATGTTGATGAGAAATTTGTTG 1368
Qy 1201 GCTACATCTACCCGCTGAGTGCCTACAGAGGGGATATCTGATCAAACTTGCTTAA 1260
Db 1369 GCTACATCTACCCGCTGAGTGCCTACAGAGGGGATATCTGATCAAACTTGCTTAA 1428

RESULT 10
ADQ24401 standard; DNA; 3998 BP.
ADQ24401;
26-AUG-2004 (first entry)
Human soft tissue sarcoma-upregulated DNA - SEQ ID 7221.
soft tissue sarcoma; cytosolic; gene therapy; vaccine; screening; human;
ds.
Homo sapiens.
WO2004048938-A2.
10-JUN-2004.
26-NOV-2003; 2003WO-US038193.
26-NOV-2002; 2002US-0429739P.
(Protein-) PROTEIN DESIGN LABS INC.
Aziz N, Ginsburg WM, Zlotnick A;
WPI; 2004-441208/41.
Early detection of soft tissue sarcoma comprises determining expression
of a gene in a first soft tissue sample and a normal soft tissue sample
and comparing the gene expression, also useful in treating soft tissue
sarcoma.
Example 2; SEQ ID NO 7221; 210pp; English.
The invention relates to a novel method for detecting soft tissue sarcoma
which comprises obtaining a first soft tissue sample from an individual
and a normal soft tissue sample from the same or different individual,
determining the expression of a gene in both samples and comparing the
expression of the gene in both soft tissue samples, where a higher level
of protein expression in the first soft tissue sample indicates the
presence of soft tissue sarcoma. The method of the invention has
cytostatic applications and may be useful for detecting soft tissue
sarcoma, possibly via gene therapy or vaccine production. The nucleic
acid sequences may be useful in diagnostic and screening applications.
The current sequence is that of a human soft tissue sarcoma-upregulated
DNA of the invention. The current sequence is not shown within the
specification per se but was submitted in CD format by the inventor.
```

```
XX
SQ Sequence 3998 BP; 1066 A; 777 C; 885 G; 1255 T; 0 U; 15 Other;
Query Match 93.3%; Score 1175; DB 12; Length 3998;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 1215; Conservative 0; Mismatches 45; Indels 3; Gaps 1;
Qy 1 ATGGGGGATCCGGCTGCGCTGCGGCGGAGCGCGCGCTTTCTGAGGCTTCGCG 60
Db 160 ATGGGGGATCCGGCTGCGCTGCGGCGGAGCGCGCGCGCTTTCTGAGGCTTCGCG 219
Qy 61 CGGCTGCGGAGTACAGAGCCGCTGCGGCTGCGCTGCGCTGCGCTGCGCTGCGCTT 120
Db 220 CGGCTGCGGAGTACAGAGCCGCTGCGGCTGCGCTGCGCTGCGCTGCGCTGCGCTT 279
Qy 121 CGGCTGCGGAGTACAGAGCCGCTGCGGCTGCGGCTGCGCTGCGCTGCGCTGCGCTT 180
Db 280 CGGCTGCGGAGTACAGAGCCGCTGCGGCTGCGGCTGCGCTGCGCTGCGCTGCGCTT 339
Qy 181 ACCACCGGGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 240
Db 340 ACCACCGGGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 399
Qy 241 GAACAGTTGCAAGGTCACAGAAACATCGTACTTGTATGAGTGTTCACATCCACTTT 300
Db 400 AAGTCTGTGTTGACTTAACCGAGAGTGACTTTGTATGAGTGTTCACATCCAGGTT 459
Qy 301 TCT---CCAAATGTCATCAGCTGTCTGTGCTTGAATCTCTGAGATGATGCTTTCG 357
Db 460 GTTCTCCAAAGTGCATACGCTGTCTGTGCTTGAATCTCTGAGATGATGCTTTCG 519
Qy 358 GAATGCTCTTATATTCAGTACACGAGGTTGTCATGAGATGATGATGATGATGATG 417
Db 520 GAATGCTCTTATATTCAGTACACGAGGTTGTCATGAGATGATGATGATGATGATG 579
Qy 418 CGAGATGTTTGGAGGCTTGTCTTTCTTCAATCAATGAGGCTATGTCATGCGGACTC 477
Db 580 CGAGATGTTTGGAGGCTTGTCTTTCTTCAATCAATGAGGCTATGTCATGCGGACTC 639
Qy 478 AAACCACTGTAATATTTGAGAGTGCAGAGATGAATGATTTTAACTCATTTGATGGA 537
Db 640 AAACCACTGTAATATTTGAGAGTGCAGAGATGAATGATTTTAACTCATTTGATGGA 699
Qy 538 CTTAGCTTCAAGAGGCAATCAGAGATGTAAGTATTTACAGACAGAGGATGCGGCT 597
Db 700 CTTAGCTTCAAGAGGCAATCAGAGATGTAAGTATTTACAGACAGAGGATGCGGCT 759
Qy 598 CCAGAGCAGATTTGCAAAATTTGCTGGCCCGAGGCTGCGCTGAGATGATCAGATGT 657
Db 760 CCAGAGCAGATTTGCAAAATTTGCTGGCCCGAGGCTGCGCTGAGATGATCAGATGT 819
Qy 658 ACTCAGCTGTTGATCTGTGAGAGCTAGAGATCATTTTAACTGGAATGTTCTCAGAAATG 717
Db 820 ACTCAGCTGTTGATCTGTGAGAGCTAGAGATCATTTTAACTGGAATGTTCTCAGAAATG 879
Qy 718 AAATGAAACATACATGATCTCAGAGATGGAAGGCAACAGTTCGTATTTAT 777
Db 880 AAATGAAACATACATGATCTCAGAGATGGAAGGCAACAGTTCGTATTTAT 939
Qy 778 CACATATTTGCGAGTAAGCAATGATGATGCGCAATTTCCAGCTTACCTTAAGAGAC 837
Db 940 CACATATTTGCGAGTAAGCAATGATGATGCGCAATTTCCAGCTTACCTTAAGAGAC 999
Qy 838 CTTATCAAAAGATGCTTATATATATCAAGACAGAAATTTCTGCTGAAATGCAATG 897
Db 1000 CTTATCAAAAGATGCTTATATATATCAAGACAGAAATTTCTGCTGAAATGCAATG 1059
Qy 898 TCGAGGCCATTTTATGATGCTTTGCGCCCTCATATTTGAAGATGATGCTGCTGCC 957
Db 1060 TCGAGGCCATTTTATGATGCTTTGCGCCCTCATATTTGAAGATGATGCTGCTGCC 1119
Qy 958 ACTCAGTCTAAGACTGCTGATGCTGATGATGATGATGATGATGATGATGATGATG 1017
```



```
QY 721 CTGAAACATACAGTCAGATCTCAGAAATGGAAGGCAAAAGTTCTGCTATTATGATCAC 780
DB 882 CTGAAACATACAGTCAGATCTCAGAGTGGAAGGCAAAAGTTCTGCTATTATGATCAT 941
QY 781 AATATTGGCCAGTAAGCAGTGGGTAATGCGCAATTCAGACCTATACCTTAAGAGACTT 840
DB 942 AATATTGGCCAGTAAGCAGTGGGTAATGCGCAATTCAGACCTATACCTTAAGAGACTT 1001
QY 841 ATCAAAAGCATGCTTCATGATGATTCACAGCAGAAAGATTCCTGCTGAAAATGGCATTTG 900
DB 1002 ATCAAAAGCATGCTTCATGATGATTCACAGCAGAAAGATTCCTGCTGAAAATGGCATTTG 1061
QY 901 AGCCCATTCCTTACATTCCTTTTGGCCCTCATATTGAAAGATTCGTCTCCACT 960
DB 1062 AGCCCATTCCTTACATTCCTTTTGGCCCTCATATTGAAAGATTCGTCTCCACT 1121
QY 961 CCAGTGTCTAAGCTGTGTAATGCTGTAATGATGATTCCTTGGGAAATGAAGAGAAAT 1020
DB 1122 CCAGTGTCTAAGCTGTGTAATGCTGTAATGATGATTCCTTGGGAAATGAAGAGAAAT 1181
QY 1021 GAAGATGTTGTAAGATGTAAGAGAGATGTCACAAAATATGACCACTGTGATCTCTA 1080
DB 1182 GAAGATGTTGTAAGATGTAAGAGAGATGTCACAAAATATGACCACTGTGATCTCTG 1241
QY 1081 CTGTGTTCCAAAGGAAATCTCTGCGACAGACAAAGTCTTTGTTGATGATGCAAAATGCTG 1140
DB 1242 CTGTGTTCCAAAGGAAATCTCTGCGACAGACAAAGTCTTTGTTGATGATGCAAAATGCTG 1301
QY 1141 GATTCCAAAGCTGCGGAGAAATTAAGTCTGAGTGAAGAGATGTTGATGAGGAAATGTTG 1200
DB 1302 GATTCCAAAGCTGCGGAGAAATTAAGTCTGAGTGAAGAGATGTTGATGAGGAAATGTTG 1361
QY 1201 GCTACATTCCTACCCGCTGAGTGTCTCAAGAGGGGATATCTGTATCAAACTTGTCTTAA 1260
DB 1362 GCTACATTCCTACCCGCTGAGTGTCTCAAGAGGGGATATCTGTATCAAACTTGTCTTAA 1421
```

RESULT 12

AAH70764

ID AAH70764 standard; cDNA; 529 BP.

AAH70764;

19-SEP-2001 (first entry)

Human cervical cancer marker nucleic acid 2038.

Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

Homo sapiens.

MO200142467-A2.

14-JUN-2001.

08-DEC-2000; 2000MO-US033312.

08-DEC-1999; 99US-0169681P.

14-MAR-2000; 2000US-0189315P.

12-MAY-2000; 2000US-0203791P.

09-JUN-2000; 2000US-0210600P.

21-JUL-2000; 2000US-0220114P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Schlegel R, Deeds J, Berger A, Zhao X;

MPI; 2001-375006/39.

New isolated nucleic acid for diagnosing and treating cervical cancer and for assessing and detecting compounds for treating the cancer.

XX Claim 1; Page 434; 1051bp; English.

PS The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy

SQ Sequence 529 BP; 156 A; 102 C; 132 G; 139 T; 0 U; 0 Other;

Query Match 36.4%; Score 458.6; DB 4; Length 529;
Best Local Similarity 97.8%; Pred. No. 3.5e-125;
Matches 486; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

QY 654 ATGTACCTCAGCTGTGATCTGTGAGCTAGAAATCAATTTACTGAAATGTTCTCAG 713

DB 22 AGTACCTCAGCTGTGATCTGTGAGCTAGAAATCAATTTACTGAAATGTTCTCAG 81

QY 714 AATGAACTGAAACATACAGTCAGATCTCAGAAATGAAGCAACAGTTCTGATTAAT 773

DB 82 AATGAACTGAAACATACAGTCAGATCTCAGAAATGAAGCAACAGTTCTGATTAAT 141

QY 774 TGATCAATATTTGCCAGTAAAGCAGTGTGAATGCCCAATTCAGCTTACCTTAAG 833

DB 142 TGATCAATATTTGCCAGTAAAGCAGTGTGAATGCCCAATTCAGCTTACCTTAAG 201

QY 834 AGACCTTATCAAAAGCATGCTTCATGATGATCAAGAGAAATTCCTGTAATGGC 893

DB 202 AGACCTTATCAAAAGCATGCTTCATGATGATCAAGAGAAATTCCTGTAATGGC 261

QY 894 ATTGTGACGCCCATTCCTTACGATTCCTTTGCCCCCTCATATTTGAAGATCTGTATGCT 953

DB 262 ATTGTGACGCCCATTCCTTACGATTCCTTTGCCCCCTCATATTTGAAGATCTGTATGCT 321

QY 954 TTCCACTCCAGTGTGAAGACTGTAATGCTGATGATGATTAATCTTGGAAATGAAGA 1013

DB 322 TTCCACTCCAGTGTGAAGACTGTAATGCTGATGATGATTAATCTTGGAAATGAAGA 381

QY 1014 GGAATATGAGA-TGTTGTGAGAAGATGTAAGA-GGAGTGCACAAAATATGACCAAGT 1071

DB 382 GGAATATGAGAAGTGTGTGAGAAGATGTAAGAAGAGAGATGTCACAAAATATGACCAAG 441

QY 1072 GTATCTCTACTGTTCCAAAGGAAATCTCTGCGACAGGACAGTCTTTGTTGATGATGCA 1131

DB 442 GTATCTCTACTGTTCCAAAGGAAATCTCTGCGACAGGACAGTCTTTGTTGATGATGCA 501

QY 1132 AATGCTGTGATTCGAA 1148

DB 502 AAGGCTGGGGATTCGAA 518

RESULT 13

AAA02535

ID AAA02535 standard; cDNA; 722 BP.

AAA02535;

19-MAY-2000 (first entry)

Human colon cancer cell line polynucleotide sequence SEQ ID NO:2526.

Human; colon cancer; tumour; diagnosis; gene expression product; probe;

detection; cancerous state; metastasis; identification; breast cancer;

estrogen receptor-negative breast cancer; lung cancer; ss.

Homo sapiens.

WO9958675-A2.

XX 18-NOV-1999.
 PD 13-MAY-1999; 99MO-US010602.
 PF 14-MAY-1998; 98US-0085426P.
 PR 15-MAY-1998; 98US-0085537P.
 PR 15-MAY-1998; 98US-0085696P.
 PR 21-OCT-1998; 98US-0105234P.
 PR 27-OCT-1998; 98US-0105877P.
 PA (CHIR) CHIRON CORP.
 PA (HYSE-) HYSEQ INC.
 XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
 PI Reinhard C, Gleese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
 PI Lamon G, Dymnac R, Ckvenjakov R, Dickson M, Dymnac S, Labat I;
 PI Lezhkowitz D, Kita D, Garcia V, Jones LM, Stachne-Crain B;
 XX WPI; 2000-126369/11.
 DR Polynucleotide library used to determine cancerous states of mammalian
 XX cells.
 PT Claim 1; Page 1020; 1097pp; English.
 PS AAA00010 to AAA02716 represent polynucleotides isolated from cDNA
 CC libraries constructed from human colon cancer cell lines. The present
 CC invention also describes a method of detecting differentially expressed
 CC genes correlated with a cancerous state of a mammalian cell, comprising
 CC detecting at least one differentially expressed gene product in a test
 CC sample derived from a cell suspected of being cancerous, where detection
 CC of the differentially expressed gene product is correlated with a
 CC cancerous state of the cell from which the test sample was derived. The
 CC polynucleotide sequences can be used in a method for detecting
 CC differentially expressed genes correlated with a cancerous state of a
 CC mammalian cell. The polynucleotides can also be used as probes for
 CC detecting and mapping related genes. They can be used in diagnosis and
 CC prognosis of diseases and disorders (e.g. identification of pre-
 CC metastatic or metastatic cancerous states, stages of cancer, or
 CC responsiveness of cancer to therapy). This is particularly for breast
 CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
 CC negative breast cancer, lung cancer, and colon cancer
 CC XX

Sequence 722 BP; 215 A; 129 C; 164 G; 193 T; 0 U; 21 Other;

Query Match 35.4%; Score 445.8; DB 3; Length 722;
 Best Local Similarity 93.0%; Pred. No. 2.7e-121;
 Matches 581; Conservative 0; Mismatches 29; Indels 15; Gaps 11;

QY 454 GAGGGCTATGTCCTCATGCGGACCTCAACGATATTTGGAGTGAGAGAA 513
 DB 1 GAGGGCTATGTCCTCATGCGGACCTCAACGATATTTGGAGTGAGAGAA 60
 QY 514 TGTTTTAACTCATTGACCTTGGACTTCAAGAGGCAATCAAGATGTAAGTAT 573
 DB 61 TGTTTTAACTCATTGACCTTGGACTTCAAGAGGCAATCAAGATGTAAGTAT 120
 QY 574 ATTACAGACAGCGGGATGCGGCTCAGAGAGCAATTTGGAAATTTGGCCGCGCT 633
 DB 121 ATTACAGACAGCGGGATGCGGCTCAGAGAGCAATTTGGAAATTTGGCCGCGCT 179
 QY 634 GGCTTCAGAGATGATCAGAGATGATGATGATGATGATGATGATGATGATGAT 693
 DB 180 GGCTTCAGAGATGATCAGAGATGATGATGATGATGATGATGATGATGATGAT 239
 QY 694 TTACTGGAATGTTCTCAGAGATGAACTGAAACATACAGTCAAGATGGAAG 753
 DB 240 TTACTGGAATGTTCTCAGAGATGAACTGAAACATACAGTCAAGATGGAAG 299
 QY 754 GGAAGAG--TTGCGCATTA--TTGATCAGATATTTGCCAGTAA--GCGATGCGTGAATGC 809
 DB 300 GGAAGAGTTTCTGCTATTATTGATCAGATATTTGCCAGTAAAGCAANTGCGATGC 359

QY 810 CGCAATTCAGGCTTATCACTTAAGAGACCTTATCAAAAGCATGCTTCAATGATCC--A 867
 DB 360 CGCAATTCAGGCTTATCACTTAAGAGACCTTATCAAAAGCATGCTTCAATGATCCAA 419
 QY 868 AGCAGAGAAATTCCTGCTGAATGGCATT--GTGACGCCATTCTTTAGC--ATTCTTTTG 925
 DB 420 GCAGAGAAATTCCTGCTGAATGGCATT--GTGACGCCATTCTTTAGC--ATTCTTTTG 479
 QY 926 CCCCTCATATTAAGATGTCGATGCTTCCCACTCCAGT--GCTAAGACTGCTGAATGTG 984
 DB 480 CCCCTCATATTAAGATGTCGATGCTTCCCACTCCAGT--GCTAAGACTGCTGAATGTG 539
 QY 985 CT--GGATGATGATTACTTGGGAATGAA--GAGGAATGAAAGATGTTGTAAGATGT- 1040
 DB 540 CTGGGATGATGATTACTTGGGAATGAAAGAGATGTTGTAAGATGTTGTAAGATGNT 599
 QY 1041 -AAAAGAGAGTGTCAAAAATATGG 1064
 DB 600 AAAAGAGAGTGTCAAAAATATTTG 624

RESULT 14
 AAH71966
 ID AAH71966 standard; cDNA; 470 BP.
 XX
 XX AAH71966;
 AC
 XX
 DT 19-SEP-2001 (first entry)
 XX
 DE Human cervical cancer marker nucleic acid 3240.
 XX
 KM Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200142467-A2.
 XX
 PD 14-JUN-2001.
 XX
 PF 08-DEC-2000; 2000MO-US033312.
 XX
 PR 08-DEC-1999; 99US-0169681P.
 PR 21-DEC-1999; 99US-0171350P.
 PR 14-MAR-2000; 2000US-0189315P.
 PR 12-MAY-2000; 2000US-0203791P.
 PR 09-JUN-2000; 2000US-0210600P.
 PR 21-JUL-2000; 2000US-0220114P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Deede J, Berger A, Zhao X;
 XX
 DR WPI; 2001-375006/39.
 XX
 PT New isolated nucleic acid for diagnosing and treating cervical cancer and
 XX for assessing and detecting compounds for treating the cancer.
 XX
 PS Claim 1; Page 636; 1051pp; English.
 XX
 CC The invention relates to novel genes (AAH68727-AAH73383) associated with
 CC cervical cancer with cytostatic activity. The nucleic acids and encoded
 CC polypeptides are useful to assess if a patient is afflicted with
 CC cervical cancer or has a pre-malignant condition; to monitor the
 CC progression of cervical cancer or a premalignant condition in a patient;
 CC and to select and/or assess the efficacy of a compound or therapy for
 CC inhibiting cervical cancer in a patient. The nucleic acids may also be
 CC useful for gene therapy
 CC XX

Sequence 470 BP; 140 A; 91 C; 114 G; 125 T; 0 U; 0 Other;
 Query Match 33.9%; Score 426.6; DB 4; Length 470;
 Best Local Similarity 98.7%; Pred. No. 1e-115;

Matches 451; Conservative 0; Mismatches 4; Indels 2; Gaps 2;
QY 654 ATGTACCTCAGCTGTGTGATCTGTGAGCCTAGGAATCATTTTACTGGAATGTTCTCAGG 713
DB 14 AGGTACCTCAGCTGTGTGATCTGTGAGCCTAGGAATCATTTTACTGGAATGTTCTCAGG 73
QY 714 AATGAACTGAAACATACAGTCTGAGATCTGAGAAATGGAAGCAAGTTCCTCATTTAT 773
DB 74 AATGAACTGAAACATACAGTCTGAGATCTGAGAAATGGAAGCAAGTTCCTCATTTAT 133
QY 774 TGATCACAATTTTGGCAGTAAAGCAGTGTGATGATGCGGAAATTCAGCCTATCAGCTAAG 833
DB 134 TGATCACAATTTTGGCAGTAAAGCAGTGTGATGATGCGGAAATTCAGCCTATCAGCTAAG 193
QY 834 AGACCTTATCAAAAGATGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 893
DB 194 AGACCTTATCAAAAGATGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 253
QY 894 ATTTGACAGCCCATTTTGTAGCATTCCTTTTGGCCCTCATATTTGAAGATCTGCTCATGCT 953
DB 254 ATTTGACAGCCCATTTTGTAGCATTCCTTTTGGCCCTCATATTTGAAGATCTGCTCATGCT 313
QY 954 TCCCACTCCAGTGTCTAAGACTGCTGAATGTGCTGAATGTATTTATCTTGGGAATGAGA 1013
DB 314 TCCCACTCCAGTGTCTAAGACTGCTGAATGTGCTGAATGTATTTATCTTGAAGATGAGA 373
QY 1014 GGAATATGAGA-TGTTGTGAGAAGATGTAAAGA-GGATGTCAAAAATATGACCAAG 1071
DB 374 GGAATATGAGA-TGTTGTGAGAAGATGTAAAGA-GGATGTCAAAAATATGACCAAG 433
QY 1072 GTATCTCTACTTGTTCAGAAAGAAATCCCTGGCAGAG 1108
DB 434 GTATCTCTACTTGTTCAGAAAGAAATCCCTGGCAGAG 470

RESULT 15
AAH69101
ID AAH69101 standard; cDNA; 461 BP.

XX AAH69101;

DT 19-SEP-2001 (first entry)

XX Human cervical cancer marker nucleic acid 375.

XX Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

OS Homo sapiens.

XX WO200142467-A2.

XX 14-JUN-2001.

XX 08-DEC-2000; 2000WO-US033312.

XX 08-DEC-1999; 99US-0169681P.

XX 21-DEC-1999; 99US-0171350P.

XX 14-MAR-2000; 2000US-0189315P.

XX 12-MAY-2000; 2000US-0203791P.

XX 09-JUN-2000; 2000US-0210600P.

XX 21-JUL-2000; 2000US-0220114P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Deeds J, Berger A, Zhao X;

XX WPI; 2001-375006/39.

PT New isolated nucleic acid for diagnosing and treating cervical cancer and
for assessing and detecting compounds for treating the cancer.

PS Claim 1; Page 168-169; 1051pp; English.

CC The invention relates to novel genes (AAH68727-AAH73383) associated with
cervical cancer with cytostatic activity. The nucleic acids and encoded
polypeptides are useful to assess if a patient is afflicted with
cervical cancer or has a pre-malignant condition; to monitor the
progression of cervical cancer or a premalignant condition in a patient;
and to select and/or assess the efficacy of a compound or therapy for
inhibiting cervical cancer in a patient. The nucleic acids may also be
useful for gene therapy

XX Sequence 461 BP; 134 A; 91 C; 109 G; 127 T; 0 U; 0 Other;

Query Match 29.8%; Score 375.4; DB 4; Length 461;

Best Local Similarity 96.2%; Pred. No. 1.6e-100;

Matches 428; Conservative 0; Mismatches 11; Indels 6; Gaps 4;

QY 654 ATGTACCTCAGCTGTGTGATCTGTGAGCCTAGGAATCATTTTACTGGAATGTTCTCAGG 713
DB 18 AGGTACCTCAGCTGTGTGATCTGTGAGCCTAGGAATCATTTTACTGGAATGTTCTCAGG 77
QY 714 AATGAACTGAAACATACAGTCTGAGATCTGAGAAATGGAAGCAAGTTCCTCATTTAT 773
DB 74 AATGAACTGAAACATACAGTCTGAGATCTGAGAAATGGAAGCAAGTTCCTCATTTAT 137
QY 774 TGATCACAATTTTGGCAGTAAAGCAGTGTGATGATGCGGAAATTCAGCCTATCAGCTAAG 833
DB 138 TGATCACAATTTTGGCAGTAAAGCAGTGTGATGATGCGGAAATTCAGCCTATCAGCTAAG 197
QY 138 TGATCACAATTTTGGCAGTAAAGCAGTGTGATGATGCGGAAATTCAGCCTATCAGCTAAG 197
DB 138 TGATCACAATTTTGGCAGTAAAGCAGTGTGATGATGCGGAAATTCAGCCTATCAGCTAAG 197
QY 834 AGACCTTATCAAAAGATGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 893
DB 198 AGACCTTATCAAAAGATGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 257
QY 894 ATTTGACAGCCCATTTTGTAGCATTCCTTTTGGCCCTCATATTTGAAGATCTGCTCATGCT 953
DB 258 ATTTGACAGCCCATTTTGTAGCATTCCTTTTGGCCCTCATATTTGAAGATCTGCTCATGCT 316
QY 954 TCCCACTCCAGTGTCTAAGACTGCTGAATGTGCTGAATGTATTTATCTTGGGAATGAGA 1013
DB 317 TCCCACTCCAGTGTCTAAGACTGCTGAATGTGCTGAATGTATTTATCTTGAAGATGAGA 376
QY 1014 GGAATATGAGA-TGTTGTGAGAAGATGTAAAGA-GGATGTCAAAAATATGACCAAG 1068
DB 377 GGAATATGAGA-TGTTGTGAGAAGATGTAAAGA-GGATGTCAAAAATATGACCAAG 436
QY 1069 GTGTATCTCTACTTGTTCAGAAAGAAATCCCTGGCAGAG 1093
DB 437 GTGTATCTCTACTTGTTCAGAAAGAAATCCCTGGCAGAG 461

Search completed: November 29, 2004, 11:45:24
Job time : 656.5 secs

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 29, 2004, 11:25:02 ; Search time 4563 Seconds
(without alignments)
4342.403 Million cell updates/sec

Title: US-10-798-532-2

Perfect score: 2206
Sequence: 1 MAGSGCAGWAEPPRFLEAFG.....VATFVPLSAVKRYLTQL 419

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

-MODE=frame+pn.model -DEV=xlp
-Q/cg2.1/USPTO.spool.p/US10798532/runat_29112004_112452_15890/app_query.fasta_1.1166
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORH=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10798532.QCGN_1.1.6317@runat_29112004_112452_15890 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl: *
1: gb_da: *
2: gb_hcg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_srs: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2198	99.6	1260	6	CQ715025 Sequence
2	2198	99.6	1260	6	AX262521 Sequence
3	2198	99.6	2598	6	AX262519 Sequence
4	2198	99.6	2909	9	HSAS36197 Homo sapi

5	2198	99.6	4065	6	CQ412057 Sequence
6	2198	99.6	4065	6	AX188296 Sequence
7	2191	99.3	1260	6	AX165543 Sequence
8	2191	99.3	1824	6	AX680136 Sequence
9	2186	99.1	1633	10	RNKISRNA
10	2186	99.1	3244	10	RNU70372
11	2185	99.0	2041	10	BC058732
12	2177	98.7	110000	2	AC129141.1
13	2160	97.9	1254	10	MMPKTS
14	1993	90.3	1367	9	BC026046
15	1893.5	85.8	2940	9	HS0805327
16	1764.5	80.0	1157	5	CR386837
17	1721	78.0	1135	9	AK058195
18	1634	74.1	2041	9	BC014917
19	1050	47.6	259562	2	AC094499
20	1045	47.4	601	10	MMKISRNA
21	908	41.2	105902	9	ALJ59699
22	908	41.2	167995	2	AC040901
23	904	41.0	211812	2	AC027205
24	805.5	36.5	211635	2	AC112763
25	789	35.8	529	6	AX186343
26	766.5	34.7	722	6	BD221413
27	729	33.0	470	6	AX187546
28	724	32.8	100207	10	AY180177
29	697.5	31.6	200095	10	AC121606
30	697.5	31.6	220806	10	AC123650
31	663	30.1	462	6	AX184680
32	369	16.7	60640	10	AY180176
33	369	16.7	185810	2	AC136052
34	361	16.4	231	6	BD220137
35	330.5	15.0	205752	2	AC119431
36	311.5	14.1	419	6	CQ393016
37	311.5	14.1	419	6	CQ399411
38	267	12.1	185810	2	AC136052
39	257	11.7	110000	8	CR382128.04
40	252	11.4	1065	10	HAMCDK2A
41	250.5	11.4	1269	10	AB096078
42	250.5	11.4	1395	10	AB096079
43	249.5	11.3	1400	10	RATCDK2A
44	249.5	11.3	2250	10	BC061832
45	248.5	11.3	2809	3	AY064464

ALIGNMENTS

RESULT 1	CQ715025	Sequence 959 from Patent W002068579.	1260 bp	DNA	linear	PAT 03-FEB-2004
LOCUS	CQ715025	Sequence 959 from Patent W002068579.	1260 bp	DNA	linear	PAT 03-FEB-2004
DEFINITION	CQ715025	Sequence 959 from Patent W002068579.	1260 bp	DNA	linear	PAT 03-FEB-2004
ACCESSION	CQ715025	Sequence 959 from Patent W002068579.	1260 bp	DNA	linear	PAT 03-FEB-2004
VERSION	CQ715025.1	GI:42275882	1260 bp	DNA	linear	PAT 03-FEB-2004
KEYWORDS						
SOURCE						
ORGANISM						
	Homo sapiens					
	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
REFERENCE						
AUTHORS	Venter, C.J., Adams, M.C., Li, P.W. and Myers, R.W.					
TITLE	Kits, such as nucleic acid arrays, comprising a majority of					
	humanexons or transcripts, for detecting expression and other uses					
	thereof					
JOURNAL	Patent: WO 02068579-A 959 06-SEP-2002;					
FEATURES						
source	Location/Qualifiers					
	1..1260					
	/organism="Homo sapiens"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:9606"					

ORIGIN

Alignment Scores: 2.81e-191 Length: 1260
Pred. No.: 2198.00 Matches: 418
Score:

Percent Similarity: 99.76% Conservative: 0
 Best Local Similarity: 99.76% Mismatches: 1
 Query Match: 99.64% Indels: 0
 DB: 6 Gaps: 0

US-10-798-532-2 (1-419) x CQ715025 (1-1260)

QY 1 MetAlaGlySerGlyCySaLaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
 Db 1 ATGGCGGGATCCGGCTGGCGTGGCGGCGGAGCCGGCGGTTTCTCGAGAGGCTTCGGG 60
 QY 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
 Db 61 CGGCTGTGGCGAGGTACAGAGCCGCTGTGGGTAGCGGCTCCCTCCGCTGGGTATCGGGTT 120
 QY 41 ArgCyCyGlyValnProGlySerProProGlyValAlaLeuGlnPheLeuProProGly 60
 Db 121 CGCTGTGGCGGCAACCTGGCTGGCCCCCGGCGCCTCAAGCAAGTTTGGCGGCAAGA 180
 QY 61 ThrThrGlyValAlaAlaSerAlaAlaGluTyrGlyPheArgGlyGluArgAlaAlaLeu 80
 Db 181 ACCACCGGGGCTGGCGCTTCTGCCCGGAGTATGTTTCCGCAAGAGAGGGCGGCTG 240
 QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
 Db 241 GAACAGTTCAGAGGTACAGAAACATCGTGAATTGTATGAGATGTTTCAATCCACTTT 300
 QY 101 SerProAsnValProSerArgCySeLeuLeuLeuGlnLeuLeuAspValSerValSerGlu 120
 Db 301 TCTCCAAATGTGCATCAACCTGTCTGTGTCTTGAACCTCGGATGTCAGTGTTCGGAA 360
 QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCySerMetTrpMetIleGlnHisCyValArg 140
 Db 361 TTGCTCTTAATTCAGAGTCCAGGGTGTTCATGTGATGATGATACAGCATTTGGCCGA 420
 QY 141 AspValLeuGlnValAlaLeuAlaPheLeuHisHisGlnGlyTyrValHisAlaAspLeu 160
 Db 421 GATGTTTGGAGGCCCTTGTCTTTCATCAATGAGGGCTATGTCATGCGGACCTTCAA 480
 QY 161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCyPheLeuLeuIleAspPheGlyLeu 180
 Db 481 CCACGTAACATATTGGAGTGCAGAAATGATTTTAACTCATTTGACCTTGGACTT 540
 QY 181 SerPheLeuGlnGlyValnGlnAspValTyrTyrIleGlnThrAspGlyTyrArgAlaPro 200
 Db 541 AGCTTCAAGAGAGCATCAGATGTAAATATTCAGACAGACGGGTATCGGGCTCCA 600
 QY 201 GluAlaGluLeuGlnAsnCySeLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCyThr 220
 Db 601 GAAGCAGAAATTCGCAAAATTCCTGGCCCAAGCTGGCTGCAGAGTATACAGATGACC 660
 QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnMetPheSerGlyMetLys 240
 Db 661 TCAAGCTGTGATCTGGAGCCCTAGGATCATTTTACTGGAAATGTTCTCAGAAATGAA 720
 QY 241 LeuLeuHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
 Db 721 CTGAATCATCATCAGATCTCAGGAATGGAAGCAACATCTCTGTATTAATGATCAC 780
 QY 261 IlePheAlaSerLysAlaValAlaAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
 Db 781 ATATTTCAGCAAGTAAAGAGGTGAATGCGCAATTCACACCTATCACCTTAAGAGACTT 840
 QY 281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCyS 300
 Db 841 ATCAAAAGCATGCTTCAATGATCCAGAGAAAGAAATCTCTGTGAAATGCAATGTGTC 900
 QY 301 SerProPhePheSerIleProPheAlaProHisIleGlnAspLeuValMetLeuProThr 320
 Db 901 AGCCCATTTCTTATGCAATCTTTTGGCCCTCATATTGAAGATCTGGTCATGCTTCCAC 960
 QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlnValnGluGluTyr 340

Db 961 CCAGTGTCAAGACTGCTGATGCTGTGATGATTAATCTTGAGAAATGAGAGAAAT 1020
 QY 341 GluAspValValGlnAspValLysGlnGluGlnCyGlnLysTyrGlyProValValSerLeu 360
 Db 1021 GAAGATGTTGTAGAAATATAAAGAGAGGTCTCAAAAATATGACCACTGATCTCTCA 1080
 QY 361 LeuValProLysGlnAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
 Db 1081 CTGTGTCCAAAGAAATCTCTGGCAGAGAGCAAGTCTTGTGTGATGATCAATGCTGCT 1140
 QY 381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
 Db 1141 GATTCCAAAGCTGCGAGAAATTAATCTGACTGGAAGAGATTTGATGGAGAGTTGTG 1200
 QY 401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
 Db 1201 GCTACATTACCCGCTGAGTGTCTTCAAGAGGAGATATCTATCAAACTTGTCTT 1257

RESULT 2
 AX262521 1260 bp DNA linear PAT 26-OCT-2001
 LOCUS Sequence 12 from Patent WO0173050.
 DEFINITION
 ACCESSION AX262521
 VERSION AX262521.1 GI:16511408
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE
 AUTHORS Meyers, R.
 TITLE 3714, 16742, 23546, and 13887 novel protein kinase molecules and
 uses thereof
 JOURNAL Patent: WO 0173050-A 12 04-OCT-2001;
 Millennium Pharmaceuticals, Inc. (US)
 FEATURES
 source 1..1260
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,81e-191 Length: 1260
 Score: 2198.00 Matches: 418
 Percent Similarity: 99.76% Conservative: 0
 Best Local Similarity: 99.76% Mismatches: 1
 Query Match: 99.64% Indels: 0
 DB: 6 Gaps: 0

US-10-798-532-2 (1-419) x AX262521 (1-1260)

QY 1 MetAlaGlySerGlyCySaLaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
 Db 1 ATGGCGGGATCCGGCTGGCGTGGCGGCGGAGCCGGCGGTTTCTCGAGAGGCTTCGGG 60
 QY 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
 Db 61 CGGCTGTGGCGAGGTACAGAGCCGCTGTGGGTAGCGGCTCCCTCCGCTGGGTATCGGGTT 120
 QY 41 ArgCyCyGlyValnProGlySerProProGlyValAlaLeuGlnPheLeuProProGly 60
 Db 121 CGCTGTGGCGGCAACCTGGCTGGCCCCCGGCGCCTCAAGCAAGTTTGGCGGCAAGA 180
 QY 61 ThrThrGlyValAlaAlaSerAlaAlaGluTyrGlyPheArgGlyGluArgAlaAlaLeu 80
 Db 181 ACCACCGGGGCTGGCGCTTCTGCCCGGAGTATGTTTCCGCAAGAGAGGGCGGCTG 240
 QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
 Db 241 GAACAGTTCAGAGGTACAGAAACATCGTGAATTGTATGAGATGTTTCAATCCACTTT 300
 QY 101 SerProAsnValProSerArgCySeLeuLeuLeuGlnLeuLeuAspValSerValSerGlu 120

REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
AUTHORS	Meyers, R.
TITLE	3714, 16742, 23546, and 13887 novel protein kinase molecules and uses therefor
JOURNAL	Patent: WO 0173050-A 10 04-OCT-2001;
FEATURES	Millennium Pharmaceuticals, Inc. (US)
SOURCE	Location/Qualifiers 1..2598 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" 268..1527 /note="named protein product" /codon_start=1 /protein_id="CAD10238.1" /db_xref="GI:16511407" /translaton="MAGSGCAGWAEPPRFLPEAFGLTWQVOSRLGSSASAYRYRCOCNPSPGALOKPLPPTGTGAASAAYGRFKERIALBOLQSHRI VTLVGFTTHFSNVSRICILRLDVSVEILLYSHQCSMMIIOHCARDYLEAFLHEGVYADLPRNLINAEENCKRLIDFGSLPKEDNDVYCIODYPAPAEIIONCIAOGSDTECTVAWDMSLGIIILEMFSGMKLHTVRSQGMKANSIAIIDITFASKAVNNALPATYLRLLIKSMIDLDPERRIPAEAMALCSPPFS.I.PFAHIEDLVWL.PPVRLNLVLDDDYCEEENEEDVDVEDVKEOCKGPVSLIVLPKENPGRQVFVEYANAGSKAAOKLTGHPMFOGKFVALTFEYPLSAVKRGVLQYTLT"
CDS	
ORIGIN	
Alignment Scores:	
Pred. No.:	7,7e-191 Length: 2598
Score:	2198.00 Matches: 418
Percent Similarity:	99.76% Conservative: 0
Best Local Similarity:	99.76% Mismatches: 1
Query Match:	99.64% Indels: 0
DB:	6 Gaps: 0
US-10-798-532-2 (1-419) x AX262519 (1-2598)	
Dn	1 MetLAaglysergLyCybaLatrpGLyAlGLubProArgphLeuGluaLaPhegLy 20
Dn	268 ATGGCGGGAAATCCGGCTCGCTGGAGCGCGGAGCGCGCGCTTTCTGAGGCCCTTCGGG 327
Dn	21 ArgLeuTrpGLInVaLGINSerArgLeuGLySerGLySerSerAlaSerValTyzArgVal 40
Dn	328 CGGCTGCGCAAGTACAGACCCTCTGGTRAGCGGCTCTCCGCTCGGTGTATCGGGTT 387
Dn	41 ArgCySGcyGLyaAnProGLySerProProGLyAlaLeuLYgLinPhLeuProProGLy 60
Dn	388 CGTGCTGCAGCAACCTCGGCTGCCCCCGGCGCCCTCAAGAGATTCTTGCGCCGACGA 447
Dn	61 ThrThrGLyAlaAlaAlaSerAlaAlaGLUtyrGLyPheAlaGLyGLuArgAlaAlaLeu 80
Dn	448 ACCACCGGGGTGGCGCTCTGCGCGCAGATGATGGTTTCCGCAAGAAGAGGGCGGCGT 507
Dn	81 GluGLInLeuGLInGLyHI;ARGAsnLI;ValThrlEuTYrGLyValPheThrLI;ehIsPhe 100
Dn	508 GAACAGTTGCGAGGCTCACAGAAACATGTCATCTTGTATGAGAGTGTTAACAATCCAATT 567
Dn	101 SerProAsnValProSerArgCYLeuLeuLeuGLuLeuLeuAspValSerValSerGU 120
Dn	568 TCTCCAAATGGCATACGCTGTCTGTGCTTGAACCTCGGAATGCAGTGTTCGGAA 627
Dn	121 LeuLeuLeuTYrSerSerHI;GLInGLyCYsSerMetTrpMetLI;eGLNIhISCYAlaArg 140
Dn	628 TTGCTCTTATTTCTACAGTACACAGAGTGTTCATATGTGATGAATAACACATTGTCCGA 687
Dn	141 AspValLeuGLuAlaLeuAlaPheLeuHI;ehIsGLyGLyValLI;hISAlaAspLeuLYs 160
Dn	688 GATTTTTTGGAGGCCCTTGTCTTTCTTATCATATAGGGCTATATGTCACAGCGACTCAAA 747
Dn	161 ProArgAsnLI;leuTYrSerAlaGLUabnGLUCysPheLYsLeuLI;asPheGLyLeu 180
Dn	748 CCAGTAGCAATTTGTGAGAGTGCAGAGAAATGAATGTTTTAAATCATTCATTCGATTCGACTT 807

```

QY 181 SerPheLeuGluGlyAangInaPValIySerTyriIeGInThraSpGlyTyraGalaPro 200
Db AGCTTCAGAGAGGACATCGATGTAAATATTATTCAGACAGCGGATTCGGCTCCA 867
QY 201 GluAlaGluLeuGlnInaPValIySerTyriIeGInThraSpGlyTyraGalaPro 220
Db GAAGGAGAAATGCAAAATGCTTGGCCAGCGCTGCGCTGCAAGATGATGCAATATACC 927
QY 221 SerAlValaPheLeuTrpSerLeuGlyIleIleLeuLeuGlnuPheSerGlyMetCys 240
Db TCGAGCTTGATCTGTGAGACCTGAGATCATTTTACTGGAATGTTCTCAGGAATGAAA 987
QY 241 LeuLeuHisThrValaIySerGlnGluTrpValaIaPheSerSerAlaIleIlePheHis 260
Db CTGAAACATGACATCGATCTCAGAAATGGAAGGCAAAACGTTCTGCTATTATTGATCAC 1047
QY 261 IlePheAlaSerTyriIeGlnInaPValIySerTyriIeGlnThraSpGlyTyraGalaPro 280
Db ATATTTGGCCAGTAAAGAGAGTGTGAATGCGCAATTCAGCCCTATCACTTAAGAGACCTT 1107
QY 281 IleLeuSerMetLeuHisAaPheProSerArgArgIleProAlaGluMetAlaLeuCys 300
Db ATTAAAAGCATGCTTCATGATGATCCAGAGAAAGAAATTCCTGTAATGCAATGTGCG 1167
QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAaPheValMetLeuProThr 320
Db AGCCCATTTCTTGTAGCATCTTGTGGCCCTCATATTGAAGATCGGTGATGCTTCCCACT 1227
QY 321 ProValLeuArgLeuLeuAaPheValIySerTyriIeGlnThraSpGlyTyraGalaPro 340
Db CCGAGCTTAAGATCTGGAATGTGCTGAGATGATGATTAATTAAGAAATGAAGGAAATAT 1287
QY 341 GluAaPheValaIyGluAaPheValIyGluGluGlyGlnIleTyriIeProValaIySerLeu 360
Db GAAGAGTGTGTGAAGATGTAAAGAGAGTGTCAAAAATATGACCAAGTGTATCTCTA 1347
QY 361 LeuValProIyGluAaPheProGlyArgGlyGlnIlePheValaIyTyriIeAaAlaGly 380
Db CTGTGTTCCAAAGGAAATCTCTGCGAGAGCAAGTCTTGTGATGCAATGCTGCTGCT 1407
QY 381 AaPheSerValaIaGlnIyLeuLeuThraGlyArgMetPheAaPheIySerPheValaIy 400
Db GATTCGAAGAGCTGCGAGAAATATGATGATGAGAGAGTGTGTGAGAGAGTGTGTG 1467
QY 401 AlaThraPheTyriProLeuSerAlaTyriIeArgGlyTyriLeuTyriGlnThraLeu 419
Db 1468 GCTACATTCACCGCTGAGTGTGCTAAGAGGAGATCTGTATCAAACTTGCCT 1524

```

RESULT 4
LOCUS HSAS36197 2909 bp mRNA linear PRI 11-JAN-2003
DEFINITION Homo sapiens mRNA for KIS protein.
ACCESSION AJ536197
VERSION AJ536197.1 GI:27657360
KEYWORDS KIS protein.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Bieche, I., Manceau, V., Curmi, P. A., Laurendon, I., Lachkar, S., Leroy, K., Vaudaud, D., Sobel, A. and Manceau, A.
AUTHORS Quantitative RT-PCR reveals a ubiquitous but preferentially neural expression of the KIS gene in rat and human
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2909)
AUTHORS Manceau, A.
JOURNAL Direct Submision
FEATURES Submitted (08-JAN-2003) Manceau A., U440, Inserm, IFM, 17 rue du fer a moulin, 75005 Paris, FRANCE
source Location/Qualifiers
1..2909
/organism="Homo sapiens"

```

/mol_type="mRNA"
/db_xref="RZPD:IMAGP98M213680"
/db_xref="taxon:9606"
/clone="I.M.A.G.E. Consortium cloneid 145548"
/tissue_type="2 pooled tumors (clear cell type)"
1..2909
/gene="KIS"
/cds
146..1405
/gene="KIS"
/contig_start=1
/product="KIS protein"
/protein_id="CAD60192.1"
/db_xref="GI:27657361"
/translation="MAGSGCAMGADPPFLAFLRQVQSRIGSSASVYRVACCG
NPGSPGALKQFLPGTTGAASAAEYGRKRALBQLOGRNIVTLVYGVTTIFSP
NYSRCLLELDIVSELLVYSHQGSMMIIOCARDVLEALFLHGEVYVADLK
PNIIIMSANEECFKIDPLSPKKNQVYIOTDGYRAPEALONCLAOGLSDTE
CTSAVDLNSLGIILEMFSKMLKRTVYSGEKANSALIDHIFSKAVVNAATRAY
LNDLIKSMILHDPSSRIPEAKLCSPPSIPAPHIEDLVMLPTVLLNLVLDYDL
ENBEYEDVDEVDKECKYGPVSLVLPKKNPGQVFEVYANAGDSKAQKLTGR
MFGKRVATFEPFLAYRGLYQTL"

```

ORIGIN

Alignment Scores:

```

Pred. No.: 9,01e-191 Length: 2909
Score: 2198.00 Matches: 418
Percent Similarity: 99.76% Conservative: 0
Best Local Similarity: 99.76% Mismatches: 1
Query Match: 99.64% Indels: 0
DB: 9 Gaps: 0

```

US-10-798-532-2 (1-419) x HSAS36197 (1-2909)

```

QY 1 MetAlaGlySerGlyCysAlaIleTrpGlyAlaGluProProArgPheLeuGlnAlaPheGly 20
Db 146 ATGGCGGAGATCCGGCTGCGCCCTGGGGCGGAGCGCCGCTTTCTGAGAGCTTCGGG 205
QY 21 ArgLeuTrpGlnValaGlnSerArgLeuGlySerGlySerSerAlaSerValTyriArgVal 40
Db 206 CGGCTGTGGCAGATACAGAGCGCTCTGGGTACCGGCTCCTCCGCTCGGTATCGGGTT 265
QY 41 ArgCysCysGlyAaPheProGlySerProProGlyValaLeuValGlnPheLeuProProGly 60
Db 266 CGCTGTGGGCAACCTGTGCTGCGCCCGCGGCTTCAACAGATCTTGTGCGGCAGGA 325
QY 61 ThrThrGlyAlaAlaAlaSerAlaIaGluTyriGlyPheArgGlyGluArgAlaIaLeu 80
Db 326 ACCACCGGGGCTGCGGCTCTGCCCGCAGATGTGTTCCGAAAGAGAGGCGCGCTG 385
QY 81 GluGlnLeuGlnGlyHisArgAanIleValIleThrLeuTyriGlyValaPheThrIleHisPhe 100
Db 386 GAACGTTGCAAGGTCACAGAAACATCGGACTTTGTATGAGTGTATCAATCCACTTT 445
QY 101 SerProAaValaProSerArgCysLeuLeuGlnuLeuAaPheValaSerValaSerGlu 120
Db 446 TCTCCAAATGTCCATCAGCTGTGCTGTGAATCTCTGATGTCACTGTTCGGGA 505
QY 121 LeuLeuLeuTyriSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
Db 506 TTGCTCTTATATTCCAGTACCAAGGTTGTTCCATGTGATATACAGATGTTGCCGA 565
QY 141 AaPheValLeuGlnAlaLeuAlaPheLeuHisGlnGlyTyriValaHisAlaAaPheLeu 160
Db 566 GATGTTTGAAGGCTTGTCTTCTTCATCATGAGGCTATGTCATGCGGAGCTCAAA 625
QY 161 ProArgAanIleLeuTrpSerAlaGluAaGluCysPheValaLeuIleAaPheGlyLeu 180
Db 626 CCAAGTAAATATTTGTGAGTGCAGAGATGATGATTTAACTATTTGACTTTGACTT 685
QY 181 SerPheLeuGluGlyAaGlnAaPheValaIySerTyriIeGlnThraSpGlyTyraGalaPro 200
Db 686 AGCTTCAGAGAGGACATCGATGTAAATATTATTCAGACAGAGCGGATTCGGCTCCA 745

```

QY 201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
 DB 746 GAAGCAGAAATTCGAAATTCCTTGGCCCGAGCTGGCTGCAGAGTATACAGAAATTCACC 805
 QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnMetPheSerGlyMetLys 240
 DB 806 TCAGCTGTTGATCTGGAGCTCAGAAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 865
 QY 241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
 DB 866 CTGAACCAATACAGTCGATCTCGAATGAGAGCAACAGTTCCTATTATGATGCAC 925
 QY 261 IlePheAlaSerLysAlaValAlaAsnAlaAlaIleProAlaTyrHisIleuArgAspLeu 280
 DB 926 ATATTGGCCAGTAAAGCAAGTGGGATGCCGCAATTCAGCCATATCACTAAGACCTT 985
 QY 281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
 DB 986 ATCAAAAGCATGCTTCATGATGATCCAGAGCAAGAAATTCCTGCAATGGCATTGTGC 1045
 QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
 DB 1046 AGCCCATCTTTAGCATTCCTTTGGCCCTCATATGGAAGATCTGGTCAATGCCACT 1105
 QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGluValAsnGluGluTyr 340
 DB 1106 CCAAGTCTAAGACTGCTGAATGCTGATGATGATTAATCTTGAGAAATGAAGGAATAT 1165
 QY 341 GluAspValValGluAspValLysGluGluCysGlnLysGlyIleProValValSerLeu 360
 DB 1166 GAAGATGTTGTAGAAATGTAAGAAAGAGATGCATCAAAATATGACACAGTGTATCTCA 1225
 QY 361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
 DB 1226 CTGTGTCGAAGGAAATCCTGCGCAGAGCAAGATCTTTGTGATGATCAAAATGCTGAT 1285
 QY 381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
 DB 1286 GATTCCAAAGCTGCGCAAAATTAATGACTGCAGAGATGTTTGATGGAGATTTGTTGG 1345
 QY 401 AlaThrPheLysProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
 DB 1346 GCTACATTCACCCGCTGAGTGCCTACAGAGGGAGATCTGTATCAAACTTGGCTT 1402
 RESULT 5
 LOCUS CQ412057 4065 bp DNA linear PAT 23-JAN-2004
 DEFINITION Sequence 19128 from Patent WO0170979.
 ACCESSION CQ412057
 VERSION CQ412057.1 GI:4119838
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Lee, J. and Lillie, J.
 TITLE Genes, compositions, kits, and method for identification,
 assessment, prevention, and therapy of ovarian cancer
 JOURNAL Patent: WO 0170979-A 19128 27-SEP-2001;
 Millennium Pharmaceuticals, Inc. (US)
 FEATURES
 source Location/Qualifiers
 1..4065
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores: 1.44e-190 Length: 4065
 Pred. No.: 2198.00 Matches: 418
 Score: 99.764 Conservative: 0
 Percent Similarity: 99.764 Mismatches: 1
 Best Local Similarity: 99.764

Query Match: 99.64% Indels: 0
 DB: 6 Gaps: 0
 US-10-798-532-2 (1-419) x CQ412057 (1-4065)
 QY 1 MetAlaGlySerGlyCysAlaIleTrpGlyAlaGluProArgPheLeuGluValPheGly 20
 DB 268 ATGGGGGAGATCCGGCTGGCGCTGGGGCGGAGCGCGCGTTTCTGGAGGCTTCGGG 327
 QY 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaTyrArgVal 40
 DB 328 CGCTGTGGCAGGTACAGAGCCGCTGGGTAGCGGCTCTCCGCTCGGTATGCGGGT 387
 QY 41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60
 DB 388 CGCTGTGGGCAACCTGTGGCTGGCCCCCGCGGCTCAACAGATCTTGTGGCGCAAG 447
 QY 61 ThrThrGlyAlaAlaIleSerAlaIleGluTyrGlyPheArgLysGluValAlaLeu 80
 DB 448 ACCACCGGGGCTGGCGCTCTGCCCGAGTATGTTCCGAAAGAGAGGGCGGCTG 507
 QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
 DB 508 GAACAGTTGACAGGTCACAGAAACATCGTGAATTTGTATGAGTGTTAACAATCCACTT 567
 QY 101 SerProAsnValProSerArgSerLeuLeuLeuGluLeuAspValSerValSerGlu 120
 DB 568 TCTCCAAATGTGCATACAGCTGTCTGTGCTTGAATCTGGATGTCAATGTTCCGGA 627
 QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
 DB 628 TTGCTCTTATATTCAGTACACAGGGTGTTCATGTGATATACAGATGTTGCCGA 687
 QY 141 AspValLeuGluValLeuValPheLeuHisGluGlyTyrValHisAlaAspLeuLys 160
 DB 688 GATGTTTGGAGGCCCTTCTTCTTCATCATGAGGCTATGTCATGCGGACCTCAAA 747
 QY 161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
 DB 748 CCAAGTACATATTTGTGAGTGCAGAGATGATATTTAACTATGATGACTTGGACTT 807
 QY 181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
 DB 808 ACCTTCAAAGAGGCAATCAGATGTAAGTATTCAGACAGAGCGGATGGGCTCA 867
 QY 201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
 DB 868 GAAGCAGAAATTCGAAATTCCTTGGCCCGAGCTGGCTGCAGAGTATACAGAAATGTACC 927
 QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnMetPheSerGlyMetLys 240
 DB 928 TCAGCTGTTGATCTGTGAGCTGAGCAATCATTTTACTGGAATGTTCTCAGGAATGAAA 987
 QY 241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
 DB 988 CTGAACCAATACAGTCGATCTCGAATGAGAGCAACAGTTCCTCATTTATGATCAC 1047
 QY 261 IlePheAlaSerLysAlaValAlaAsnAlaAlaIleProAlaTyrHisIleuArgAspLeu 280
 DB 1048 ATATTGGCCAGTAAAGCAAGTGGTGAATGCCGCAATTCAGCCTATCACTAAGACCTT 1107
 QY 281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
 DB 1108 ATCAAAAGCATGCTTCATGATGATCCAGCAAGAAATTCCTGCAATGGCATTGTGC 1167
 QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
 DB 1168 AGCCCATCTTTAGCATTCCTTTGGCCCTCATATGGAAGATCTGTATGCTTCCACT 1227
 QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluTyr 340
 DB 1228 CCAAGTCTAAGACTGCTGAATGCTGATGATGATTAATCTGGAAGAAAGAGGAATAT 1287

QY 341 GIUAAPVAlValGlUAAPVAllysglUGlUCyGInlySTyrglYProVaAlaSerleu 360
|
|
|
Db 1288 GAAGATGTTGTAAGATGTAAAGAGGATGTCAAAAATATAGACAGATGTCTTA 1347
|
|
|
QY 361 LeuValProlYsglUaAPProglYArGlyGlnValPheValGluTYrAlaAaAlaGly 380
|
|
|
Db 1348 CTTGTTCCAAAGGAAATCTGTGACAGAGCAAGTCTTGTGATGATGCAATGCTGAT 1407
|
|
|
QY 381 AaPSeRlyAlaAlaGlnlySleuLeuThrglYArMeCpheaPglYlySPhaVal 400
|
|
|
Db 1408 GATTCCAAAGCTCCGAGAAATTAATCTGAGAGATGTTTGAATGAGAGATTGTGTG 1467
|
|
|
QY 401 AlAthPheTYrProLeuSerAlaTYrlySArGlyTYrleuTYrGlnThrleuLeu 419
|
|
|
Db 1468 GCTACATTCTACCGCTGAGTGTCTTACAGAGGGATATCTGTATCAAACTTGCTT 1524
|
|
|
RESULT 6
AX188296 4065 bp DNA linear PAT 06-AUG-2001
LOCUS Sequence 3991 from Patent WO0142467.
DEFINITION AX188296
ACCESSION AX188296
VERSION AX188296.1 GI:15139769
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Schlegel, R., Deede, J., Berger, A. and Zhao, X.
Auteurs Genes, compositions, kits, and methods for identification,
titles assessment, prevention, and therapy of cervical cancer
Patent: WO 0142467-A 3991-14-JUN-2001;
JOURNAL Millennium Predictive Medicine, Inc. (US)
FEATURES
source Location/Qualifiers
1.4065
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 1,44e-190 Length: 4065
Score: 2198.00 Matches: 418
Percent Similarity: 99.76% Conservative: 0
Best Local Similarity: 99.76% Mismatches: 1
Query Match: 99.64% Indels: 0
DB: Gaps: 0
US-10-798-532-2 (1-419) x AX188296 (1-4065)
QY 1 MetAlAGlySerGlyCySaAlaTpglYAlaGluProProArpPheLeuGlnAlaPheGly 20
|
|
|
Db 268 ATGGCGGGATCCGGCTGCGCTGGGGCGGAGCCGCGGCTTTTTCGAGGCTTCGGG 327
|
|
|
QY 21 ArGleuTYrGlnValGlnSerArGleuGlySerGlySerSerAlaSerValTYrArGVal 40
|
|
|
Db 328 CCGCTGTGCGAGGTACAGAGCCCTGAGGTAAGCGGCTCCCTCCGCTGCGGTATCGGGTT 387
|
|
|
QY 41 ArGcyCyGglYAsnProglYSerProProglYAlaLeuLyglInPheLeuPProProGly 60
|
|
|
Db 388 CGTGTCTGGGGCAACCTGTGCTGCCCCCGGCGCCCTCAAGCAATTTTGGCCGAGGA 447
|
|
|
QY 61 ThrThGlyAlaAlaAlaSerAlaAlaGluTYrGlyPheArGlySglUaArGAlaAlaLeu 80
|
|
|
Db 448 ACCACCGGGGCTGCGGCTTGTGCGCGGATATGTTTCCGCAAAAGAGGGGGCGCTG 507
|
|
|
QY 81 GlUglnLeuGlnGlyHISArGAsnIleValThrleuTYrGlyValPheThrIleHisPhe 100
|
|
|
Db 508 GAACAGTTGCAAGGTACAGAAACATCTGTGATGAGAGTGTTCACATCCACTTT 567
|
|
|
QY 101 SerProArValProSerArGcySleuLeuLeuGlnleuAspValSerValSerGlu 120
|
|
|
Db 568 TCTCCAAATGTGCATCAAGCTGTGTGTGTGAACCTCTGTGATGTCAAGTGTTCGAA 627
|
|
|

QY 121 LeuLeuLeuTYrSerSerHISglnglYCySerMetTTPMeCIIeGlnHISCySaAlaArG 140
|
|
|
Db 628 TTGCTCTTATATATCCAGTACACAGGGTTTCTTCATGTGGATATACAGATTGTCCCGA 687
|
|
|
QY 141 AaPValLeuGlnAlaLeuAlaPheLeuHISISglUGlyTYrValHISAlaAPleuLyS 160
|
|
|
Db 688 GATGTTTGGAGCCCTTGCTTTCTTCATCTAGAGGGCTATGTCTCATCGGACCTCAA 747
|
|
|
QY 161 ProArGAsnIleLeuThrSerAlaGluAsnGluCySPhaLeuIleAPheGlyLeu 180
|
|
|
Db 748 CACGTAACATATGTGAGTGCAGAGATGATGTTTAACTATTCATTCATTCGACTT 807
|
|
|
QY 181 SerPheLySglYAsnGlnaAPVAllySTyrlIeGlnThrAPglTYrArGAlaPro 200
|
|
|
Db 808 AGCTTCMAAGAAAGGCAATCAGATGTAAAGTATATTCAGACAGAGGGATCGGCTCCA 867
|
|
|
QY 201 GlUAlaGlnLeuGlnAsnCylLeuAlaGlnAlaGlyLeuGlnSerAPThrglUCyThr 220
|
|
|
Db 868 GAAGCAGATTCCAAAATTTGCTTGCCCGGCTGCTGCAGAGTATCAAGAAATGTACC 927
|
|
|
QY 221 SerAlaValAPleuThrSerleuGlyIleIleLeuLeuGluMetPheSerGlyMetLyS 240
|
|
|
Db 928 TCAAGCTGTGATCTGTGAGCTAGAGATCATTTTACGAAATGTTTCAGAAATGAA 987
|
|
|
QY 241 LeuLyHISThrValArSerglnglUTpLyAlaAsnSerSerAlaIleIleAPHIS 260
|
|
|
Db 988 CTGAACATACAGTCAATCAGATCTCAGAAATGAAAGCAAGTTCGTATTATTCATCAC 1047
|
|
|
QY 261 IlePheAlaSerlyAlaValAlaAlaIleProAlaTYrHISleuArGpleu 280
|
|
|
Db 1048 AATATTTGCCAGTAAGACAGTGTGATGCGCAATTCAGCTTACCTTAAGAACCTT 1107
|
|
|
QY 281 IleLySerMetLeuHISaAPProSerArGAlaIleProAlaGluMeAlaLeuCyS 300
|
|
|
Db 1108 ATCAAAAGCATCTTCATATATATCAACAGAAATTCCTGCTGAATGGCATTTGTC 1167
|
|
|
QY 301 SerProPhePheSerIleProPheAlaProHISISglUaAPleuValMetLeuProThr 320
|
|
|
Db 1168 AGCCCATCTTTAGCATTCCTTTTGCCCTCATATTAAGATCTGTGATGCTCCACT 1227
|
|
|
QY 321 ProValLeuArGleuLeuAsnValleuAPaAPArTYrleuGlyAsnGluGluTYr 340
|
|
|
Db 1228 CCAGTCTTAAGCTGTGATGTCTGTGATGTATCTTGAGAAATGAAGAAATAT 1287
|
|
|
QY 341 GIUAAPVAlValGlUAAPVAllysglUGlUCyGInlySTyrglYProVaAlaSerleu 360
|
|
|
Db 1288 GAAGATGTTGTAAGATGTAAAGAGGATGTCAAAANTATGACACAGTGTATCTCTA 1347
|
|
|
QY 361 LeuValProlYsglUaAPProglYArGlyGlnValPheValGluTYrAlaAaAlaGly 380
|
|
|
Db 1348 CTTGTTCCAAAGGAAATCTGTGACAGAGCAAGTCTTGTGATGATGCAATGCTGAT 1407
|
|
|
QY 381 AaPSeRlyAlaAlaGlnlySleuLeuThrglYArMeCpheaPglYlySPhaVal 400
|
|
|
Db 1408 GATTCCAAAGCTCCGAGAAATTAATCTGAGAGATGTTTGAATGAGAGATTGTGTG 1467
|
|
|
QY 401 AlAthPheTYrProLeuSerAlaTYrlySArGlyTYrleuTYrGlnThrleuLeu 419
|
|
|
Db 1468 GCTACATTCTACCGCTGAGTGTCTTACAGAGGGATATCTGTATCAAACTTGCTT 1524
|
|
|
RESULT 7
AX166543 1260 bp DNA linear PAT 22-JUN-2001
LOCUS Sequence 34 from Patent WO0138503.
DEFINITION AX166543
ACCESSION AX166543
VERSION AX166543.1 GI:14546888
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1

AUTHORS Ploewman, G.D., Whyte, D., Manning, G.S., Sudarsanam, S.S., Martinez, R., Flanagan, P. and Clary, D.S.
 TITLE Novel human protein kinases and protein kinase-like enzymes
 JOURNAL Patent: WO 0138503-A 34 31-MAY-2001;
 Sugen, Inc. (US)
 FEATURES Location/Qualifiers
 source 1..1260
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,236-190 Length: 1260
 Score: 2191.00 Matches: 417
 Percent Similarity: 99.52% Conservative: 0
 Best Local Similarity: 99.52% Mismatches: 2
 Query Match: 99.32% Indels: 0
 DB: 6 Gaps: 0
 US-10-798-532-2 (1-419) x AX166543 (1-1260)
 QY 1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
 DB 1 ATGGCGGGATCCGGCTGCGCTGGGGCGCGAGCCCGCGTTTCTGGAGGCTTCGGG 60
 QY 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
 DB 61 CGGCTGTGGCAGGTACAGAGCCGCTGGGTAGCGGCTCCCGCCCTGGGTATCCGGGTT 120
 QY 41 ArgCysCysGlyValanProGlySerProProGlyValLeuGlyGlnPheLeuProProGly 60
 DB 121 CGGTGCTGGGGCAACCTGGCTGCGCCCGCGGCGCTCAAGCATGTTGGCGGCAAGA 180
 QY 61 ThrThrGlyValAlaAlaSerAlaAlaGluTyrGlyPheArgGlyGluArgAlaAlaLeu 80
 DB 181 ACCACCGGGGCTGCGGCTCTGCGCGAGTATGTTCCGCAAGAGGGGCGGCTG 240
 QY 81 GluGlnLeuGlnGlyHisArgAniLeuValThrLeuTyrGlyValPheThrIleHisPhe 100
 DB 241 GAACAGTTGCAGGGTCACAGAAACATGTCATTTGATGAGTGTTCACATCCACTTT 300
 QY 101 SerProAnValProSerArgCysLeuLeuLeuGluLeuLeuAnValSerValSerGlu 120
 DB 301 TCCTCAAATGTGCTCATCAGCTGTGCTTGAACCTCGTGAATGTCAGTGTTCGGA 360
 QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
 DB 361 TTGCTCTTATATTCAGTCCAGGGTGTTCATGTGATGATACAGCATTTGCCCCA 420
 QY 141 AspValLeuGlnAlaLeuAlaPheLeuHisHisGlnGlyTyrValHisAlaAspLeuLys 160
 DB 421 GATGTTTGGAGGCCCTTGCTTTTTCATCATGAGGGCTATGTCCATGCGGACCTCAA 480
 QY 161 ProArgAniLeuLeuTrpSerAlaGluAnGluCysPheLeuLeuLeuPheGlyLeu 180
 DB 481 CCAAGTAACATATGTGGAGTGCAGAAATGAATGTTTAACTCATGACTTGTGA 540
 QY 181 SerPheLeuGlnGlyAnGlnAnPheValTyrTrpIleGlnThrAspGlyTyrArgAlaPro 200
 DB 541 AGCTTCAAGAAAGGCATCGAGTGTAAATATATTCAGACACAGCGGTATCCGGCTCA 600
 QY 201 GluAlaGluLeuGlnAnGlyAnGlyAnGlyAnGlyAnSerAspThrGluCysThr 220
 DB 601 GAAGCAGAAATGCAAAATTCCTTGCCCAAGGCTGCTGCAGAGTGTACAGAAATGACC 660
 QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluLeuPheSerGlyMetLys 240
 DB 661 TCAAGCTGTATCTGTGGAGCTTGAAGATCATTTTACGTGAATGTTCTCAGAAATGAAA 720
 QY 241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAnSerSerAlaIleIleAspHis 260
 DB 721 CTGAACATATCAGTCAATCTCAGGAATGAGGCAAAACAGTTCTGTATATTATGATCAC 780

QY 261 IlePheAlaSerValAlaValAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
 DB 781 ATATTGTCAGTAAGACATGTGTGATGCGCAATTCAGCTATACCTTAAGAACCTT 840
 QY 281 IleLysSerMetLeuHisAspAspProSerArgGlyIleProAlaGluMetAlaLeuCys 300
 DB 841 ATCAAAAGCATCTTCATGATATCCAGCAGAAATATTCCTGCTGAATGCAATGTGTC 900
 QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
 DB 901 AGCCCATCTTATGATTCCTTTGCCCCCTCATATGAAGATCTGTGATCTCCCACT 960
 QY 321 ProValLeuArgLeuLeuAnValLeuAspAspAspTyrLeuGlyAnGlnGluGlyTyr 340
 DB 961 CCAGGTCAAGCTGCTGAATGTGTGATGATATTCCTTGAGAAATGAAAGAAATAT 1020
 QY 341 GluAspValAlaGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
 DB 1021 GAAGATGTTGTGAAGATTAAGAGAGATCTCAAAAATATGACACAGTGTATCTCTA 1080
 QY 361 LeuValProLysGluAspProGlyValArgGlyGlnValPheValGluTyrAlaAnAlaGly 380
 DB 1081 CTGTGTCCAAAGGAATCTCTGCAGAGGACAAGTCTTGTGATGATGCAATGCTGGT 1140
 QY 381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
 DB 1141 GATTCCAAAGCTGCGCAGAAATTAAGTACTGATGAGAGATTTGATGGAGAGTTGTGG 1200
 QY 401 AlaThrPheTyrProLeuSerAlaTyrLysAspGlyTyrLeuTyrGlnThrLeuLeu 419
 DB 1201 GCTACATTTACCCGCTGAGTGCCTTCAAGAGGAGATATCTTATCAACCTTGCTT 1257
 RESULT 8
 AX680136 1824 bp DNA linear PAT 29-MAR-2003
 LOCUS AX680136
 DEFINITION Sequence 21 from Patent WO0181555.
 ACCESSION AX680136
 VERSION AX680136.1 GI:29369916
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 Gururajan, R., Paterson, C., Recipon, S.A., Tribouley, C.M.,
 Hafalia, A., Khan, F.Y., Yue, H., Au-Young, J., Bandman, O.,
 Baughn, M.R., Borowsky, M.L., Burford, N., Burrill, J.D., Elliot, V.S.,
 Gandhi, A.R., Kearney, L., Lal, P., Lu, D.A., Lu, Y., Tang, Y.T.,
 Zingler, K.A., Griffin, J.A., Hillman, D.J., Marcus, G.A., Nguyen, D.B.,
 Policky, J., Ramkumar, J., Thornton, M., Walla, N.K. and Walsh, R.T.
 Human kinases
 Patent: WO 0181555-A 21 01-NOV-2001;
 Incyte Genomics, Inc. (US)
 FEATURES Location/Qualifiers
 source 1..1824
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 3107952CB1"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,056-190 Length: 1824
 Score: 2191.00 Matches: 417
 Percent Similarity: 99.52% Conservative: 0
 Best Local Similarity: 99.52% Mismatches: 2
 Query Match: 99.32% Indels: 0
 DB: 6 Gaps: 0
 US-10-798-532-2 (1-419) x AX680136 (1-1824)
 QY 1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20

```

Db      169 ATGGCCGGATCCCGCTGCGCGCGGAGCGCGCGCGCTTTCTGGAGCGCTTCGGG 228
Qy      21 ArgLeuThrGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
Db      229 CGGCTGGAGAGGTACAGACCGCTCGGGAGCGGCTCCCGCTCGGGGATCGGGATT 288
Qy      41 ArgCysArgGlyAsnProGlySerProProGlyAlaLeuGlyGlnPheLeuProProGly 60
Db      289 CGCTGCTGGGGGACCGCTGGCTCGCCCGCGCGCTCAAGAGATTCTTGGCGGACGGA 348
Qy      61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgGlyGluArgAlaAlaLeu 80
Db      349 ACCACCGGGGCTCGCGCTCTGCGCGGAGTATGTTTCGCAAAAGAGGGGCGGCTG 408
Qy      81 GlnGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
Db      409 GAAACAGTTGCAGGGGTACAGAAACATCGTACCTTTGTATGAGGTGTTTACAAATCCACTT 468
Qy      101 SerProAsnValProSerArgCysLeuLeuLeuGlnLeuLeuAspValSerValSerGlu 120
Db      469 TCTCCAAATGTGCATACAGCTGCTGTTGCTTGAACCTCGTGATGTCAGTGTTCGGAA 528
Qy      121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysValAlaArg 140
Db      529 TTGCTCTTATATTCACAGTCCAGCGGTGTTCATGTGATGATACAGCATTTGTGCCCA 588
Qy      141 AspValLeuGlnValaLeuAlaPheLeuHisHisGlnGlyTyrValHisAlaAspLeuVal 160
Db      589 GATGTTTGGAGGCGCTTGTCTTTCTTCATCATGAGGGCTATGTCATGCGGACCTCAAA 648
Qy      161 ProArgAsnIleLeuTrpSerAlaGluAsnGlyCysPheLeuLeuIleAspPheGlyLeu 180
Db      649 CCAAGTAACATATATGTGGAGTCCAGAGAAATGATTTAACTCATTTGACTTGGACTT 708
Qy      181 SerPheGlyGlnGlyAsnGlnAspValValTyrIleGlnThrAspGlyTyrArgAlaPro 200
Db      709 AGCTTAAAGAGAGGCAATCAGATGTAAAGTAAATATTCAGACACGCGATATCGGGCTCCA 768
Qy      201 GlnAlaGlnLeuGlnAsnGlyLeuAlaGlnAlaGlyLeuGlnSerAspThrGlyCysThr 220
Db      769 GAAAGCAATTCGAAATTCCTGCGCCAGCGCTGCGCTGCAAGTGTACAGAAATGACC 828
Qy      221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnLeuTrpPheSerGlyMetLys 240
Db      829 TCAGCTGTGATCTGTGGAGCTTAGAATCATTTTACTGGAATGTTTCCAGGAATGAAA 888
Qy      241 LeuLeuHisThrValArgSerGlnGluTrpValAlaAsnSerSerAlaIleIleAspHis 260
Db      889 CTGAAACATACAGCTCAGATCTCAGAAATGAGAAACAGTCTCTATTTATTTGATCAC 948
Qy      261 IlePheAlaSerIleValaValaAlaAlaIleProAlaTyrHisIleuArgAspLeu 280
Db      949 ATATTTCAGCAATTAACAGTGTGAATGCGCAATTCAGCTTACCTTACAGAGACCTT 1008
Qy      281 IleLeuSerMetLeuHisAspAspProSerArgArgIleProAlaGlnMetAlaLeuCys 300
Db      1009 ATCAAAAGCATGCTTCATGATGATCCAAAGAGAAATTCCTGCTGAATGGCATTGTGC 1068
Qy      301 SerProPhePheSerIleProPheAlaProHisIleGlnAspLeuValMetLeuProThr 320
Db      1069 AGCCCAATCTTTTAGCATTCCTTTTGGCCCTCATATGAAGATGTGCTATGCCACT 1128
Qy      321 ProValLeuArgLeuLeuAsnValaLeuAspAspAspTyrLeuGlnGlnGlnGlyTyr 340
Db      1129 CCAATCTTAAGATCTGCTGAATGCTCGATGATGATTAATTTAGAAATGAAGAGAAATAT 1188
Qy      341 GlnAspValaValaGlnAspValaValGlnGlnGlyCysGlnLeuTyrGlyProValaValSerLeu 360
Db      1189 GAAAGATGTTGTAGAAAGATGTAAAGAGAGAGTCAAAAATATGAGACAGTGGATCTCTA 1248
Qy      361 LeuValProLeuGlnAsnProGlyArgGlyGlnValaPheValaGlnTyrAlaAsnAlaGly 380

```

```

Db      1249 CTTGTTCCAAAGGAAATCTGCGAGAGACAAATCTTTGTTAGTANGCAAAATCTGCT 1308
Qy      381 AspSerIleAlaAlaGlnLeuLeuThrGlyArgMetPheAspGlyLysPheValaVal 400
Db      1309 GATTCCAAGCTGCGGAGAAATTAATTAATGATGAGAGATGTTTGAAGGAAATTTGTGG 1368
Qy      401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
Db      1369 GCTACATTTACCCGCTGAGTCTCTAAGAGGAGATATCTGTATCAAACTTGTCTT 1425

```

RESULT 9

```

LOCUS      RNKISRNA      1633 bp      mRNA      linear      ROD 08-JUL-2002
DEFINITION R.norvegicus mRNA for KIS protein.
ACCESSION  X98374
VERSION    X98374.1 GI:1403531
KEYWORDS   kis gene.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.

```

REFERENCE

```

1 Maucuer, A., Ozon, S., Manceau, V., Gavet, O., Lawler, S., Curmi, P. and
  Sobel, A.
  KIS is a protein kinase with an RNA recognition motif
  JOURNAL   U. Biol. Chem. 272 (37), 23151-23156 (1997)
MEDLINE    97435279
PUMED      9287318

```

REFERENCE

```

2 (bases 1 to 1633)
Maucuer, A.
Direct Submision
Submitted (10-JUN-1996) A. Maucuer, INSERM U440, 17 rue du Fer
Moulin, 75005 Paris, FRANCE

```

FEATURES

```

source      location/Qualifiers
            1..1633
            /organism="Rattus norvegicus"
            /mol_type="mRNA"
            /db_xref="taxon:10116"
            /cissue_type="brain"
            /clone_lib="7"
            /dev_stage="neonate"
            1..1633
            /gene="KIS"
            /gene="KIS"
            /gene="KIS"
            /codon_start=1
            /protein_id="CAA67021.1"
            /db_xref="GI:1403532"
            /db_xref="GOA:O63285"
            /db_xref="SwIsa-Prot:O63285"
            /trna_label="MAGSGCAGGAPPRFLEAFGLMQVQSRILSGSSASVYRVACCG
            TPSPGALKOFLPGTTGAASAAEYGRKRALBLOQGRNVTLYGVFTTHFSP
            NPSRCILLETIVSESLIVSSHQGSMNIOHCARDVLEALFHLHEGVHADLK
            PNIIISASBNECFKLIDFGLSKFKENQVYKIQTGTCYARPELONCLAQAGSDTE
            CTSADVLSLGIILLEMFSGMLKRTVSSQEKANSSAIIDHIFASKAVVNAATPAIH
            LKDLIKSMIHDEPSRRIPAEVALCSPPFSIPAPRIEDLVMLPTVVLNLNLDDYL
            ENBEDVEVEDVEKECCQKYGVSILVPEKNPGRGVFEVYANAGDSKAAQKLLTGR
            MFDGKFAVATFPFLAYKRGVLYGTL"

```

gene

```

CDS

```

```

ORIGIN
Alignment Scores:
Pred. No.:      5,05e-190      Length:      1633
Score:          2186,00      Matches:      415
Percent Similarity: 99,52%      Conservative: 2
Best Local Similarity: 99,05%      Mismatches:  2
Query Match:    99,09%      Indels:      0
DB:              10      Gaps:      0
US-10-798-532-2 (1-419) x RNKISRNA (1-1633)

```

```

Qy      1 MetAlaGlySerGlyCysAlaIleTrpGlyAlaGluProProArgPheLeuGlnAlaPheGly 20

```


Db 39 ATGCCCGGGTCCGGCTGGCGCTGGGGCGCCGAGCCGCCGCTTCTCGAGAGCTTCGGG 98
 Qy 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
 Db 99 CGCGTGTGGCAGGTAACAGAGCCGCCCTTAGGACGCGGCTCCCGCTCGGTATACCGGGG 158
 Qy 41 ArgGlySerGlyAspProGlySerProProGlyValAlaLeuGlnPheLeuProProGly 60
 Db 159 CGCTGCTGGCGACTCAGAGCTCGCCCCCGGGCGCCCTCAAGAGATTCCTGCTCCGGGA 218
 Qy 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgGlyGlnArgAlaAlaLeu 80
 Db 219 ACCACCGGGGCTGGCGCTCGGTGGGAGTATGTTTCCAAAAGAGGGAGCGCTG 278
 Qy 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
 Db 279 GAGCAGTTTCAGGGTCACAGAAACATCGACTTATACGAGTCTTATCCATACACTTC 338
 Qy 101 SerProAsnValProSerArgCysLeuLeuLeuGlnLeuLeuAspValSerValSerGlu 120
 Db 339 TCTCCAAATGTGCATCAGCTGTCTGTGTGAATCTCTGAGTTCAGCGTTTCGGA 398
 Qy 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
 Db 399 TTGCTTGTATATTCCAGTCAACAGGTTGCTCATGTGATGATCCAGCACTGTGCAGA 458
 Qy 141 AspValLeuGlnAlaLeuAlaPheLeuHisHisGlnGlyTyrValHisAlaAspLeuLys 160
 Db 459 GATGTCTGAGAGCCCTTGCTTTTCTTCAACATGAGGGCTATGCTCATGACAGCTCAAG 518
 Qy 161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLeuLeuLeuAspPheGlyLeu 180
 Db 519 CACGAAACATCTGTGAGTGGGAGAACGATGCTTAAAGCTTATGACTTGAAGTTC 578
 Qy 181 SerPheLeuGlnGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
 Db 579 AGCTTCAAGAGAGCAATCAGAGCTGGAATATTCAGACGACGGGTATCGAGCTCT 638
 Qy 201 GluAlaGlnLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
 Db 639 GAAGCAGAACTGCAGAAATGCTTGCCCAAGCGCGCTGACAGATACAGAGTAC 698
 Qy 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuGlnMetPheSerGlyMetLys 240
 Db 699 TCAAGCTGTGATCTGAGAGCTTGAATCATTTTACTGAAATGTTCTCAGAAAGAA 758
 Qy 241 LeuLysHisThrValArgSerGlnGluTyrLysAlaAsnSerSerAlaIleIleAspHis 260
 Db 759 CTGAAACATACAGTCAAGATTCAGAGTGAAGGCAACAGTTCTGCTATATTGATCAT 818
 Qy 261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
 Db 819 ATATTGGCCAGTAAACAGAGTGAATGCGCAATTCACGCTATCACCACAGACCTT 878
 Qy 281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGlnMetAlaLeuCys 300
 Db 879 ATCAAAAGCATGCTTATGACGACCAAGCAGAGAGATCCCTCTGAGAGGCTGTGTC 938
 Qy 301 SerProPhePheSerIleProPheAlaProHisIleGlnAspLeuValMetLeuProThr 320
 Db 939 AGCCCATTTCTTAGCATCTTTTGTGCCCTCATATTGAAGATCTGGATGATCTTCAACT 998
 Qy 321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGlnGluTyr 340
 Db 999 CCAAGTCTCAGACTCTCCCAATGTGCTGATGATGATCATCTTGAATAATGAAATGATAT 1058
 Qy 341 GluAspValValGlnAspValLysGlnGluCysGlnLysTyrGlyProValValSerLeu 360
 Db 1059 GAAGAGTGTGTAAAGATGTTAAAGAGAGATGTTCAGAAATATGACAGAGGTTTCTCTG 1118
 Qy 361 LeuValProLysGlnAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
 Db 1119 CTGTGTTCCAAAGAAATCTCTGCGACAGAGCAAGTCTTTGTTAGATACGAAACGCTGGA 1178

Qy 381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
 Db 1179 GATTCCAAAGCTGCTCAGAAATGCTGACTGGAGAGATTTGAGAGGAGATTTGTGTG 1238
 Qy 401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
 Db 1239 GCTACATTCTACCCCTGATGCTCTACAGAGGAGATTCCTTATCAAACTTGCTT 1295
 RESULT 10
 RNU70372 3244 bp mRNA linear ROD 08-SEP-1999
 LOCUS Rattus norvegicus PAM COOH-terminal interactor protein 2 mRNA,
 DEFINITION complete cds.
 ACCESSION U70372
 VERSION U70372.2 GI:5821767
 KEYWORDS
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 3244)
 Alam,M.R., Caldwell,B.D., Johnson,R.C., Darlington,D.N., Maine,R.E.
 and Eipper,B.A.
 Novel proteins that interact with the COOH-terminal cytosolic
 routing determinants of an integral membrane peptide-processing
 enzyme
 J. Biol. Chem. 271 (45), 28636-28640 (1996)
 JOURNAL MEDLINE 97067094
 PUBMED 8910496
 REFERENCE 2 (bases 1 to 3244)
 Caldwell,B.D., Darlington,D.N., Penzes,P., Johnson,R.C.,
 Eipper,B.A. and Maine,R.E.
 The Novel Kinase P-CIP2 Interacts with the Cytosolic Routing
 Determinants of the Peptide Processing Enzyme Peptidylglycine
 alpha-Amidating Monooxygenase
 J. Biol. Chem. 274 (1999) In press
 JOURNAL 4 (bases 1 to 3244)
 REFERENCE 3 (bases 1 to 3244)
 Alam,R., Caldwell,B.D., Johnson,R.C., Darlington,D.N., Maine,R.E.
 and Eipper,B.A.
 Direct Subunit
 Submitted (11-SEP-1996) Neuroscience, Johns Hopkins University
 School of Medicine, 725 N. Wolfe Street, WBSB 902A, Baltimore, MD
 21205, USA
 JOURNAL 4 (bases 1 to 3244)
 REFERENCE 4 (bases 1 to 3244)
 Caldwell,B.D., Darlington,D.N., Penzes,P., Johnson,R.C.,
 Eipper,B.A. and Maine,R.E.
 Direct Subunit
 Submitted (03-SEP-1999) Neuroscience, Johns Hopkins University
 School of Medicine, 725 N. Wolfe Street, WBSB 902A, Baltimore, MD
 21205, USA
 REMARK
 COMMENT Sequence update by subunitter
 On Sep 3, 1999 this sequence version replaced gi:1698778.
 FEATURES
 source
 1..3244
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 162..1421
 /function="serine protein kinase"
 /note="P-CIP2"
 /codon_start=1
 /product="PAM COOH-terminal interactor protein 2"
 /protein_id="AAC53031.2"
 /db_xref="GI:5821768"
 /translation="MAGSGCAGAPPRFLAFRLMVOYSLGSSSVYRVCCG
 7PGSPGALAKOFLPCTGCAASAEYGRERALLPOLGHRNVTLYGVTFHSP
 NTPSKLLLELDVSVSELYTSSHOGCSMMITQCARVDLEALAFHHBGTVDHDK
 PNNTLSAENECFKLIDGLSPKGNQDQKTYIQDGYRPALELDQNCIAQAGQSTG
 CTSADVLSIGITILEMFSGMKLTAKTYSQEWKANSALIIDHIFASKAVANNAIPAYH
 LKDLIKSMLEDDPSRRIPALMLCSPPFSIPAPATIEDLVMLPTVLRILNVLDDYL
 ENDEVEDVEVKEBCQKYGPVVSLVLPKEMPGRGQYFEVYANNGDSKAAQKLTGR

ORIGIN
MEDKRFVATFPYPLSAVKRGYLYQTL"

Alignment Scores:

Pred. No.:	1,31e-189	Length:	3244
Score:	2186.00	Matches:	415
Percent Similarity:	99.52%	Conservative:	2
Best Local Similarity:	99.05%	Mismatches:	2
Query Match:	99.09%	Indels:	0
DB:	10	Gaps:	0

US-10-798-532-2 (1-419) x RNUT0372 (1-3244)

QY 1 MetAlaGlySerGlyCyValATPrgIyAlaGluProProAArgPheLeuGluAlaPheGly 20

DB 162 ATGGCCGGGTCGGCTGCGCGTGGGGGCGCGAGCCGCGCTTCCTCGAGGCGCTTCGGG 221

QY 21 ArgLeuTrpGlnValGlnSerAlaGleuGlySerGlySerSerAlaSerValTYrAqVal 40

DB 222 CGGCTGTGGAGGTACAGAGCCGCGCTTAGCGAGCGGCTTCGCGCTCGGTACCGGCTG 281

QY 41 ArgCySeGlyValaNPProGlySerProProGlyValaLeuGlyGlnPheLeuProProGly 60

DB 282 CGCTGCTGGGCACTCCAGGCTGCGCCCGCGGCGCTTCAGAGAGTTCTGCTCGGGA 341

QY 61 ThrTrpGlyValAlaAlaSerAlaAlaGluTYrGlyPheArgGlyAlaAlaAlaLeu 80

DB 342 ACCACCGGGGCTCGCGCTCGCGCTGCGAGTATGTTCCGAAAGAGAGGCGAGCGCTG 401

QY 81 GlnGlnLeuGlnGlnGlnAlaPheLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 100

DB 402 GAGCAGATGTGAGGGGTACAGAGAACATGTGACTTTATACGAGATCTTTACATACACTTC 461

QY 101 SerProAsnValProSerArgCyValLeuLeuGlnLeuLeuLeuValSerValSerGly 120

DB 462 TTCGCCAATGTGCAATCAAGCTGCTGCTTCTTGAATCCTGAGTGTCAAGCTTCGGA 521

QY 121 LeuLeuLeuTYrSerSerHisGlnGlyCySerMetTrpMetIleGlnHisCyValaArg 140

DB 522 TTGCTGTATATTCAGTCAACAGGAGGTGCTTCATGTGATTCAGCACTGTGCAGA 581

QY 141 ArgValLeuGlnAlaLeuAlaPheLeuHisGlnGlyTYrValHisAlaAlaPheLeu 160

DB 582 GAGTCTGTGAGGCGCTGCTTCTTCTTCAACATGAGGGCTATATCCATGAGACCTTCAG 641

QY 161 ProArgAsnIleLeuTrpSerArgAlaGluAsnGluCyPheLeuLeuLeuLeuLeu 180

DB 642 CCAAGAAACATCTCTGAGAGTGGCGAGAACAGAGTCTTAAAGCTTATGACTTGAATC 701

QY 181 SerPheLeuGlnGlnGlnAlaPheValTYrTYrIleGlnThrAspGlyTYrAqValPro 200

DB 702 AGCTTCAAGAGAGCATCAGAGCATGAGATATATTCAGACAGACGGGTATCCAGCTCCT 761

QY 201 GlnAlaGlnLeuGlnHisGlnValAlaGlnAlaGlnGlnGlnGlnGlnGlnGlnGlnGln 220

DB 762 GAAAGCAAGATGAGATATCTTGGCCGAGCGCGCTGAGAGTGTATACAGAGTACC 821

QY 221 SerAlaValaPheLeuTrpSerLeuGlyIleIleLeuLeuGlnLeuMetPheSerGlyMetIys 240

DB 822 TCGAGTGTGATCTCTGAGACCTTGGAGATCTTAACTGGAATATGTTCTCAGAGAAAGAA 881

QY 241 LeuLeuHisThrValArgSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 260

DB 882 CTGAAACATGTCAGTCAATCTCAGAGTGTGAGAGCAAAAGCTTCTGCTATATATGATCAT 941

QY 261 IlePheAlaSerIysValaValaValaAlaAlaIleProAlaTYrHisLeuAqValPro 280

DB 942 ATATTGTCAGTAAAGAGTGTGAATGCGCAATTCACGCTTATACCTTCAAGAGACCTT 1001

QY 281 IleLeuSerMetLeuHisAspArgProSerArgArgIleProAlaGlnMetAlaLeuCyS 300

DB 1002 ATCAAAAGCATGCTTCAATGACGACCAAGAGAGAGATCCCTCTGATGATGCGCTTGTGC 1061

QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAlaPheValMetLeuProThr 320

DB 1062 AGCCCATCTCTTAAAGCATCTCTTGTGCGCTTCATATGAAGATCGGTGTGCTTCCAACT 1121

QY 321 ProValLeuArgLeuLeuAsnValLeuAspArgTYrLeuGlyValaGlnGlnGlnGlnGln 340

DB 1122 CCAGTCTCAAGCTCTCAATGTGCGAGTGTGAGATATCTTGAATATGAAGATGAATAT 1181

QY 341 GluAspValaValaGlnAspValIlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 360

DB 1182 GAAAGATGTTGTGAAGATCTTAAAGAGAGTCTCAGAAATATGAGACCACTGCTTCTCG 1241

QY 361 LeuValProGlyGlnAspProGlyArgGlyGlnValPheValGlnTYrAlaAsnAlaGly 380

DB 1242 CTGTGTCCAAGAGAAATCTGCGAGAGACCAAGCTTGTGTGAGTACCAACCTGGA 1301

QY 381 AspSerValaAlaGlnIlyLeuLeuTrpGlyArgMetPheAspGlyPhePheVala 400

DB 1302 GATTCCAAAGCTGCTCAGAAATGCTGACTGGAGAGATTTGACGAGAACTTTGTTGG 1361

QY 401 AlaThrPheTYrProLeuSerAlaTYrIlyArgGlyTYrLeuTYrGlnThrLeuLeu 419

DB 1362 GCTACATTCTACCGCTGAGTGCCTTACAGAGGGATATCTTATCAAACTTGCTT 1418

RESULT 11

BC058732 2041 bp mRNA linear ROD 30-JUN-2004

LOCUS Mus musculus kinase interacting with leukemia-associated gene (leukemin), mRNA (cDNA clone MGC:64770 IMAGE:6414877), complete cds

ACCESSION BC058732

VERSION BC058732.1 GI:37194892

KEYWORDS MGC.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 2041)

AUTHORS Strausberg, R.D., Fellngold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shemmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Datchenko, L., Marusina, K., Farmer, A.F., Rudin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uedl, T.B., Toshiyuki, S., Carinci, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullan, S.J., Bosak, S.A., McMan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A.C., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmitz, J., Myers, R.M., Butcherfield, Y.S., Kizylinski, M.I., Skaleks, U., Small, D.E., Scherker, A., Schein, J.E., Jones, S.J., and Matra, M.A.

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

PUBMED 12477932

REFERENCE 2 (bases 1 to 2041)

AUTHORS Strausberg, R.

TITLE Direct Submision

JOURNAL Submitted (24-SEP-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

CONTACT: MGC help desk

EMAIL: cgabs-r@mail.nih.gov

CDNA Library Preparation: Dr. Jim Lin, University of Iowa

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LUNL)

DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant.
 Web site: <http://genome.uiowa.edu/tom-casavant@uiowa.edu>
 Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
 Bonaldi, M.F., Akabogu, I., Bair, T., Bair, J., Crouch, K., Davis, A., Fishler, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K., Scheetz, T., Smith, C., Snir, E., Tack, D., Trout, K., Walters, J., Casavant, T., Soares, M.B.

Clone distribution: MCC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
 Series: Plate: Row: Column: 0
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6754447.

FEATURES

Source
 1..2041
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="MGC:64770 IMAGE:641487"
 /issue_type="Brain, pool of enriched mouse brain tissues from 13.5, 14.5, 16.5 and 17.5 dpc"
 /clone_lib="NIH BMAP_FRO"
 /lab_host="DH10B"
 /note="Vector: pYX-ASC"
 1..2041
 /gene="Kist"
 /note="synonym: KIS"
 /db_xref="LocusID:16589"
 /db_xref="MGI:1341908"
 154..1413
 /gene="Kist"
 /codon_start=1
 /product="kinase interacting with leukemia-associated gene (statmin)"
 /protein_id="AAH58732.1"
 /db_xref="GI:37194893"
 /db_xref="MGI:1341908"
 /db_xref="LocusID:16589"
 /translation="MAGSGCAGWGPREFLEAFGRMLQVQSRGSSSVYRVRCCG
 TPGSPGALKQFLPPTGTGAASAAEYGRKERALBOLQGHNIYTLGVTFHSP
 NVPSRCLLELDVSVSELLYSHOGCSMMMOHCARDVLEALAHGEGYHADLK
 PNIIISAENCKKIDPLSPKSNODVXYIOTDPTRAPEALQICLAQAGIADST
 CTSADVNLGILILEMPSGKMKHTVRSQEMKANSATIDHFAKAYVNAIPIYH
 LBDLKSMLHDDPGRIPALMALCSPFSIPAPHIEDLVMLPTVLRLLNLDVLY
 ENDEYEDVVEDVKECKXGVPVSLVPEKNPGRQVFEVYANAGDSKAAQKLTR
 MEDGKRVATFYPLSAVKRGYLYQTLL"

CDS

gene

1..2041
 /gene="Kist"
 /note="synonym: KIS"
 /db_xref="LocusID:16589"
 /db_xref="MGI:1341908"
 /db_xref="LocusID:16589"
 /translation="MAGSGCAGWGPREFLEAFGRMLQVQSRGSSSVYRVRCCG
 TPGSPGALKQFLPPTGTGAASAAEYGRKERALBOLQGHNIYTLGVTFHSP
 NVPSRCLLELDVSVSELLYSHOGCSMMMOHCARDVLEALAHGEGYHADLK
 PNIIISAENCKKIDPLSPKSNODVXYIOTDPTRAPEALQICLAQAGIADST
 CTSADVNLGILILEMPSGKMKHTVRSQEMKANSATIDHFAKAYVNAIPIYH
 LBDLKSMLHDDPGRIPALMALCSPFSIPAPHIEDLVMLPTVLRLLNLDVLY
 ENDEYEDVVEDVKECKXGVPVSLVPEKNPGRQVFEVYANAGDSKAAQKLTR
 MEDGKRVATFYPLSAVKRGYLYQTLL"

ORIGIN

Alignment Scores:
 Pred. No.: 8.5e-190 Length: 2041
 Score: 2185.00 Matches: 415
 Percent Similarity: 99.28% Conservative: 1
 Best Local Similarity: 99.05% Mismatches: 3
 Query Match: 99.05% Gaps: 0
 DB: 10

US-10-798-532-2 (1-419) x BC058732 (1-2041)

QY 1 MetAlAG1SerG1yCyba1aTtpG1yAlaG1uProProArpPheUenG1uAlaPheG1y 20
 Db 154 ATGGCGGGGTCGGCGTCGGCGGGGCGCGAGCCGCGGCTTCTGAGGCGCTTCGGG 213
 QY 21 ArgLeuTrpGlnValGlnSerArgLeuG1ySerG1ySerSerAlaSerVal1yArgVal 40
 Db 214 CGGCTGTGGGAGCTCCAGAGCCGCTGGGCAAGCGCTCGTGGCTTCGGGTACCGGGTG 273
 QY 41 ArgCyCyG1yAnpProG1ySerProProG1yAlaLeuLysGlnPheLeuProProG1y 60
 Db 274 CGGTGTGGCGGTACCCCGGGCTGCGCCCGGGCCCTCAAGAGTTCTCGCTCCGGGA 333
 QY 61 ThrThrG1yAlaAlaAlaSerAlaAlaG1uTyG1yPheArgLysG1uArgAlaAlaLeu 80

Db 334 ACCACCGGGGCTCGCGCTCGCGCGGAGTATGTTCCGCAAGAGGCGCGCTG 393
 QY 81 GluGlnLeuGlnG1yH1ArgAn1leValThrLeuTyG1yValPheThr1leH1Spre 100
 Db 394 GAGCAGTTGACAGGTCACAGGAACCTCGACTTATATGAGTCTTATCACTACCTTC 453
 QY 101 SerProAnValProSerArgCySerLeuLeuG1uLeuLeuAspValSerValSerG1u 120
 Db 454 TCTCCCATGTCCATCATCGCTGTCTGCTTGAATCTCTGAGTCTCATGTTTCGGA 513
 QY 121 LeuLeuLeuTySerSerSerH1sg1nG1yCySerMetTrpMet1leGlnH1sCysAlaArg 140
 Db 514 TTGCTTTTATATTCAGTCATCAGGGCTGCTCAGTGGATATATCAGCAGCTGCGCA 573
 QY 141 AspValLeuG1uAlaLeuAlaPheLeuH1S1sg1uG1yTyArgValH1SAlaAspLeuLys 160
 Db 574 GATGTTCTGGAGGCGCTCGCTTCTTCCACATGAGGCGTACATGCACAGCTCAAA 633
 QY 161 ProArgAn1leLeuTrpSerAlaG1uAnG1uCyPheLysLeu1leAspPheG1yLeu 180
 Db 634 CACGAAATCTCTGTGAGTCGAGAAATGAGTGTATTAAGCTTATTAAGCTTGAATTC 693
 QY 181 SerPheLysG1uG1yAsnG1nAspValLysTyR1leGlnThrAspG1yTyArgAlaPro 200
 Db 694 AGCTTCAAGAGGCAATCAGAGCGTAAAGTATATTCAGACAGAGCGGTATCGCGCTCT 753
 QY 201 GluAlaG1uLeuGlnAnCySerLeuAlaG1uAlaG1yLeuGlnSerAspThrG1uCyThr 220
 Db 754 GAAGCAGAGCTCAGAACTGCTTGGCCGAGCGCGCTGCAGAGATGATCAGAAATGATC 813
 QY 221 SerAlaValAspLeuTrpSerLeuG1y1leLeuLeuG1uMetPheSerG1yMetLys 240
 Db 814 TCAAGTGTATCTGTGAGCTCGAGATCATTTATCGAAATGTTCTTCAAGAAATGAG 873
 QY 241 LeuLysH1S1ThrValArgSerGlnG1uTrpLysAlaAsnSerSerAla1le1leAspH1S 260
 Db 874 CTGAAGAACATCAGTCAGATCTCAGGATGGAAGGAAACAGTTCGCTATTTATGATCAT 933
 QY 934 ATATTTGCGCAGTAAGACGATGTGATGCGCAATTCAGCTTATCACTCAGAGCTT 993
 Db 281 1leLysMetLeuH1SAspAspProSerArgAlaG1u1leProAlaG1uMetAlaLeuCyS 300
 QY 994 ATCAAAAGCATCTCATGATGACCCGCGCAGAAAGATCCCTGAGATGCGATTTGTC 1053
 QY 301 SerProPhePheSer1leProPheAlaProH1S1leG1uAspLeuValMetLeuProThr 320
 Db 1054 AGCCATTTCTTATGATCTCTTTGCGCTCATATTTGAAGATCTGTGATGCTTCGACT 1113
 QY 321 ProValLeuArgLeuLeuAnValLeuAspAspAspTyRLeuG1yAsnG1uG1uTyR 340
 Db 1114 CCAGTGTCTCAGGCTCTCATGTGCTGATGATGATTAAGTCTTGAAGAAATGAAGATAT 1173
 QY 341 GluAspValValG1uAspValLysG1uG1uCySerGlnTyG1yProValValSerLeu 360
 Db 1174 GAAGATGTTGAGAGATGTGAAGAGAGTCTCAGAAATATGACCACTGTGTTCTCG 1233
 QY 361 LeuValProLysG1uAnpProG1yArgG1yGlnValPheValG1uTyRAlaAn1leG1y 380
 Db 1234 CTGTTCCAAAGAAATCTCTGCGAGAGACAAAGTCTTGATGATGAGTACCGCAAGCTGTG 1293
 QY 381 AspSerLysAlaAlaGlnLysLeuLeuThrG1yArgMetPheAspG1yLysPheValVal 400
 Db 1294 GATTCCAAAGCTCTCAGAAATGTTGCTGACTGGAGAGATGTTGACGCGGAAGTTGTTGG 1353
 QY 401 AlaThrPheTyRProLeuSerAlaTyR1yAspG1yTyRLeuTyG1ThrLeuLeu 419
 Db 1354 GCTACATCTACCGCTGAGTGTCTTACAGAGAGGATCTTATCAAACTTGTCTT 1410
 RESULT 12
 AC129141_1/c

MPCOMMENT

Sequence split into 4 fragments LOCUS AC129141 Accession AC129141

Fragment Name	Begin	End
AC129141_0	1	110000
AC129141_1	100001	210000
AC129141_2	200001	310000
AC129141_3	300001	363206

Continuation (2 of 4) of AC129141 from base 100001 (AC129141 Rattus norvegicus clone CH2)

Alignment Scores:

Pred. No.:	1.18e-186	Length:	110000
Score:	2177.00	Matches:	414
Percent Similarity:	99.28%	Conservative:	2
Best Local Similarity:	98.81%	Mismatches:	3
Query Match:	98.69%	Indels:	0
DB:	2	Gaps:	0

US-10-798-532-2 (1-419) x AC129141_1 (1-110000)

QY 1 MetAlaGlySerGlyCysAlaATrGlyAlaGluProProAArgPheLeuGluAlaPheGly 20

DB 70316 ATGGCCGGGCTCCGCTGGCGGTGGGGCCGAGCCGCCGCTTCCCGAGGCCCTCGGG 70257

QY 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40

DB 70256 CCGCTGGCGAGGTACAGAGCCGCTTAGGCGAGCGGCTCCGCGCTCGGTGTACCGGGGTG 70197

QY 41 ArgCysCysGlyAsnProGlySerProProGlyValLeuLeuArgGlnPheLeuProProGly 60

DB 70196 CGCTGCTGGCGAGCTCCAGGCTCGCCGCCGCGCCCTCAAGAGCTTCCGCTCCGGGA 70137

QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgGlyGluArgAlaAlaLeu 80

DB 70136 ACCACGGGGCTGCCGCTCGCTGGCGAGTATGTTCCGAAAGAGAGCGGCGCTG 70077

QY 81 GluGlnLeuGlnGlnGlnHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100

DB 70076 GAGCAGCTTGGAGGTCACAGGAACATCGTACTTTATACGAGCTCTTACCATACACTTC 70017

QY 101 SerProAsnValProSerArgCysLeuLeuLeuGlnLeuLeuAspValSerValSerGlu 120

DB 70016 TCTCCAAAGTGGCATTACCGCTGCTGTGCTTGAACCTCTGAGATGACGTTTGGAA 69957

QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140

DB 69956 TTGCTTGTATATTCACGTACCAAGGTTGCTCATGTGATGATACAGCATCTGGCAA 69897

QY 141 AspValLeuGlnValAlaLeuAlaPheLeuHisGlnGlyTyrValHisAlaAspLeuLys 160

DB 69896 GATGTTCTGGAGGCCCTTGTCTTTCTTTCACCATAGAGGCTATGTCATGACAGCTCAA 69837

QY 161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLeuLeuLeaAspPheGlyLeu 180

DB 69836 CCAGCAAACTCTGTGAGTGGAGAGAGAGAGAGGCTTTAAGCTTATGACTTGGACTC 69777

QY 181 SerPheLeuGlnGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200

DB 69776 AGTTTCAAGAGAGCATACAGACGTGAATATTTCAGACAAACGGGTATCGAGCTCC 69717

QY 201 GluAlaGlnLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGlyCysThr 220

DB 69716 GAAGCGAATCTGCAGAACTGCTTGGCCCAAGCGGCTCGACAGATACAGAGTGTACC 69657

QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnLeuPheSerGlyMetLys 240

DB 69656 TCGAGCTGTGATCTTGGAGCCTTGGAAATATTTCAGAAATGTTCTCAGAAATGAAA 69597

QY 241 LeuLysHisThrValArgSerGlnGluTyrTrpLysAlaAsnSerSerAlaIleIleAspHis 260

DB 69596 CTGAACATATCATCATGATCTCAGAGTGAAGCAACAGTTCTGTATATTATGATCAT 69537

QY 261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280

DB 69536 ATATTTCAGTAAGACAGTGTGAATGCCGAATTCAGCCTTACCTCAGAGACTT 69477

QY 281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300

DB 69476 ATCAAAAGCATGCTTCTATACACACCAAGCAAGAGATCCCTGTGAGATGSCCTTGTC 69417

QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320

DB 69416 AGCCCATCTTATGACTTCTTTTGCCCTCATATTGAAGATCTGTGATGCTTCCAAT 69357

QY 321 ProValLeuAlaArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGlnGluGly 340

DB 69356 CCAAGCTCAGACTCTCTCAATGTGCTGATGATGATCTTGAATAATGAATGAATAT 69297

QY 341 GluAspValValGluAspValLysGlnGluCysGlnLysTyrGlyProValValSerLeu 360

DB 69296 GAAAGTGTGTGAAGATTTAAAGAGAGTGTCAAGAAATAGACCACTGTTTCTCG 69237

QY 361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380

DB 69236 CTGTTCCAAAGAAAATCTCTGGCAGAGCAAGTCTTGTGAGTACCAACCAACCTGGA 69177

QY 381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400

DB 69176 GATTCCAAAGCTGCTCAGAAATTTGCTGACTGGAGAGATTTGATGGGAAGTTGTG 69117

QY 401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419

DB 69116 GCTACATTACCCGCTGAGTGTCTTACAGAGAGGAGATCTTATCAACACTTGTCT 69060

RESULT 13

MMPKIS 1254 bp mRNA linear ROD 08-JUL-2002

LOCUS Mus musculus partial mRNA for KIS protein kinase (Kis gene).

DEFINITION Y10725

ACCESSION Y10725.2 GI:21726712

VERSION Kis gene; protein kinase.

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1254) Maucuer, A., Gavet, O., Lawler, S., Curmi, P. and Sobel, A.

AUTHORS J. Biol. Chem. 272 (37), 23151-23156 (1997)

MEDLINE 97435279

PUBMED 9287318

REFERENCE 2

AUTHORS Maucuer, A.

TITLE Direct Submision

JOURNAL Submitted (23-JAN-1997) Maucuer A., INSERM U440, U440, 17 Rue du Fer Moulin, Paris, 75005, FRANCE

REMARK 3 (bases 1 to 1254) revised by (4)

COMMENT On Jul 10, 2002 this sequence version replaced gi:1806129. Overlaps with related sequences X62320, X10725.

FEATURES

source

1..1254

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="K22"

/issue_type="brain"

/clone_lib="lambda gt10"

/dev_stage="neonate"

1..1254

/gene="Kis"

11..>1254

CDS

```

/gene="Kis"
/codon_start=1
/product="Kis protein kinase"
/protein_id="CAA7171.2"
/db_xref="GI:21726713"
/db_xref="GOA:P97343"
/db_xref="Swiss-Prot:P97343"
/tranrelat="MAGSCGACGAPRFLAFCRLMOVOSRLGSSASVYRVACCG
TGESPPGALKOFLPREGTGAASABYFRKRKRALBQJQHRNVTLYGVTFTHFSP
NPSRCLLELDVSVSELLYSHQSCSMNMIQHCARDVLAFLHHEGVHADLK
NPNIISANECEFKLIDFGLSPKEGNQDVKYIOTDGYRABEALONCIAQJQSDTE
CSAVDMISGIIILEMFSKMLKHTFRQEWKANSASAIIDHIFASKAVNVAIYAH
LMDLKSMHDDPGRRIIPALMALCSPPSIPAPPHIDLVMLPTVLRILNVLDDYL
EMEDERYEDVBEKCECKKQVSVLIVKENPGRQGVVEYANNQDSAAQKLTGR
MFDGRFVATFPPLSAHYKGYL"

```

ORIGIN

Alignment Scores:

Pred. No.:	8,34e-188	Length:	1254
Score:	2160.00	Matches:	410
Percent Similarity:	99.28%	Conservative:	1
Best Local Similarity:	99.03%	Mismatches:	3
Query Match:	97.91%	Indels:	0
DB:	10	Gaps:	0

US-10-798-532-2 (1-419) x MMPKIS (1-1254)

```

QY 1 MetAlaGlySerGlyCysAlaArgGlyAlaGluProArgPheLeuGluAlaPheGly 20
DB 11 ATGGCGGGGTCGGCGCTCGCGTGGCGCGCGAGCCGCCGCTTCTGAGAGCGCTTCGGG 70
QY 21 ArgLeuThrGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
DB 71 CGAGCTGTGACAGGTCAGAGCGCGCTGGCGAGCGCTGCTGCGCTGAGTACCGGGTGG 130
QY 41 ArgCysCysGlyAsnProGlySerProGlyAlaLeuGlyGlnPheLeuProGly 60
DB 131 CGGTGTGGGCTACCCCGGGCTCGCCCCCGGCGCTTCAAGCATTCCTGCTCCGGGA 190
QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgGlyGluArgAlaLeu 80
DB 191 ACCACCGGGGCTCGCGCTCGCGCGGAGTATGTTCCGCAAGAAGAGGGCGCGCTG 250
QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
DB 251 GAGCAGTTTCAGAGGTCACAGAAACATGACTTATATAGAGTCTTTTACATACACTTC 310
QY 101 SerProAsnValProSerArgCysLeuLeuLeuGlnLeuLeuAspValSerValSerGlu 120
DB 311 TCTCCCAATGTGCATCAGCTGTCTGCTGCTTGAACCTCTGATGTCAGTGTTCGGA 370
QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTyrMetIleGlnHisCysAlaArg 140
DB 371 TTGCTTTTATATTCAGTATCATCGAGGCTGCTCATGTGATGATACAGCATGTGCGCA 430
QY 141 AspValLeuGlnAlaLeuAlaPheLeuHisGlnGlyTyrValHisAlaAspLeuVal 160
DB 431 GATGTTCTGAGGCTCTCGCTTTTCTTCAACATGAGGCTTACCTCATGACAGACCTCAA 490
QY 161 ProArgAsnIleLeuTyrSerAlaGluAsnGluCysPheLeuLeuIleAspPheGlyLeu 180
DB 491 CCAGCAAAACATCTCTGAGAGTGGGAGATGATGTTTAACTTATTAAGCTTTTGAAC 550
QY 181 SerPheLeuGluGlyAsnGluAspValIleValTyrIleGlnThrAspGlyTyrArgAlaPro 200
DB 551 AGCTTCAAGAAGAGCATCAGACGTAAGATATATTCAGACAGACGGATATCGCGCTCT 610
QY 201 GluAlaGluLeuGlnAsnGlyLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
DB 611 GAGACGAGCTGACAGACTGCTTGGCCAGGCGCGCTGAGAGTATTCAGATATATACC 670
QY 221 SerAlaValAspLeuTyrSerLeuGlyIleIleLeuLeuGlnMetPheSerGlyMetCys 240

```

```

DB 671 TCAGCTGTGATCTGTGAGACCTCGAATCATTTTACTGAAATGTCTCAGAAATCAAG 730
QY 241 LeuValHisGlnValArgSerGlnGluTyrPylAlaAsnSerSerAlaIleIleHisHis 260
DB 731 CTGAAGCATACAGTACGATCTCAGAAAGGAGGCAACAGTTCGCTATTAATGATCAT 790
QY 261 IlePheAlaSerValAlaValAlaAlaIlePheAlaTyrHisIleLeuArgAspLeu 280
DB 791 ATATTTCGATGAAAGCATGTGTGATGATCCGCAATTCAGCTTATCATCTCAGACACT 850
QY 281 IleLeuSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
DB 851 ATCAAAAGCATGCTTCATGATGACCCGCGAGAAATCCCTGCTGAGATGGCATTTGGC 910
QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
DB 911 AGCCCATCTTTCAGATTCCTTTTGCCTCCCTCATATATGAAATGTGTATGCTTCGACT 970
QY 321 ProValLeuArgLeuLeuAsnValIleuAspAspAspTyrLeuGlyAsnGluGluTyr 340
DB 971 CAGTGTCTCAGGCTCTCTCAATGTCTGATGATGATTTACCTTGAAATGAAATGATAT 1030
QY 341 GluAspValValGluAspValIleGluGluGluCysGlnTyrGlyProValIleSerLeu 360
DB 1031 GAAAGATGTTGTGAGAGATGTGAAAGAGAGTGTCAAAATATGACACAGTGTCTCTG 1090
QY 361 LeuValProValGluAsnProGlyArgGlyGlnValIlePheValGluTyrAlaAsnAlaGly 380
DB 1091 CTGTGTCAMAGAAATCTGTGAGAGACAAAGTCTGTGATGATGCGCAACCTGCT 1150
QY 381 AspSerValAlaAlaGlnLeuLeuLeuThrGlyArgMetPheAspGlyIlePheValVal 400
DB 1151 GATTCCAAGCTGCTCAGAGTGTCTGACTGGAGAGATGTTGACGGAAGTTTTGTG 1210
QY 401 AlaThrPheTyrProLeuSerAlaTyrIleArgGlyTyrLeu 414
DB 1211 GCTACATTTTACCCGCTGAGTGTCTTCAAGAGGGATATCTT 1252

```

RESULT 14
BC026046
LOCUS
DEFINITION
Hom sapiens (human)
ACCESSION
BC026046
VERSION
BC026046.1 GI:19684094
KEYWORDS
MGC.
SOURCE
Hom sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1367)
Struhsberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,F.S., Shemen,C.M., Schuler,G.D.,
Alichaul,S.F., Zeeberg,B., Buetow,K.H., Scheffer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Mex,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Uebelin,T.B., Toshilviki,S.,
Carinci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullaby,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gamarale,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Halyk,S.W.,
Villalob,D.K., Wazny,D.M., Sodergren,E.D., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Keltman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Boutford,G.G., Blakesley,R.W., Touchwood,J., Green,B.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,T.S., Krzyzanski,M.I., Skalska,U., Smalins,D.B.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL

PUBMED 12477932
 REFERENCE 2 (bases 1 to 1367)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (20-MAR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgabs-rc@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada
 Info@cgsc.bc.ca
 Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Odi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Lilia Prabhu, Parvaneh Saeedi, JR Santos, Angeline Scherch, Ursula Skalska, Duane Smalins, Jeff Scott, Miranda Tsai, George Yang, Jacque Schein, Asim Siddiqui, Rob Holt, Marco Marra.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAL Plate: 42 Row: f Column: 6
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21389438.
 Location/Qualifiers
 1. 1367
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:39391 IMAGE:4649261"
 /cissue_type="Eye, retinoblastoma"
 /clone_lib="NIH MGC_16"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 1. 1367
 /gene="KIS"
 /db_xref="LocusID:127933"
 159. 1193
 /gene="KIS"
 /codon_start=1
 /product="kinase interacting statminin"
 /protein_id="AAH26046.1"
 /db_xref="GI:19684095"
 /transcript="MAGSCAMGABPREFLEAFGRMLVOYSLGSSASVYRVRCG NGSPRGALKOPLPQGTGAASAAEYGRKEAALDLOGRNITVLYGVPTIHSP NVPSCRLLLELDVSYSEILYSHOGCSMMNIOHCAKVLEALATLHHEGVYADLK PKRILWASNECEKLTDFGLSFKENQDVKTQTDYRAPEBLQCLQAGLOSTPE CSAVDLWSLGIILFEMFGMKLKHTRVQEWKANSATIDHIFASKAVNNAIPAYH LBDLIKSMLEDDPSRRIPEAMALCSPFSPFIPAPAHIEDLVMLPTVLRILNLYDDYL EMBEEYEGC"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.77e-172 Length: 1367
 Score: 1993.00 Matches: 386
 Percent Similarity: 92.14% Conservative: 1
 Best Local Similarity: 91.90% Mismatches: 2
 Query Match: 90.34% Indels: 31
 DB: 9 Gaps: 1
 US-10-798-532-2 (1-419) x BC026046 (1-1367)
 QY 1 MGCALGLYSerGLYcysAATATPGIYAAAGLUPROProARphereuGLUAAaheGLY 20
 DB 159 ATGGCGGAGATCCGGCTGGCGCTGGAGCGCGCCGCGGTTTCTGAGAGGCTTCGGG 218

QY 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
 DB CGGCTGTGGCAGATACAGAGCCGCTGGGGTAGCGGCTCTCCGCTCGGATGAGGGTT 278
 QY 41 ArgCysCysGlyAsnProGlySerProProGlyValAlaLeuValGlnPheLeuProProGly 60
 DB CGCTGTGGCGGACCTGTGGCTGGCCCCCGCGCGCTCCACAGCAATTCCTTGGCCGACGA 338
 QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgValGluArgAlaAlaLeu 80
 DB ACCACCGGGGCTGGCGCTCTGGCCCGCAGATGAGTTTCCGAAAGAGGGCGGCTG 398
 QY 81 GlnGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
 DB GAACAGTTGACAGGTCACAGAAACATCGGACTTTGTATGAGTGCTTTACATCACTT 458
 QY 101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
 DB TCTCCAAATGTGCCATACAGCTGTCTGTGCTTGAACCTCGAGATGACATGTTTCGGA 518
 QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
 DB TTGCTCTTATATTCAGATCACCAGGCTTCTTCATGTGATATACAGCATTTGTCCCGA 578
 QY 141 AspValLeuGluAlaLeuAlaPheLeuHisHisGlnGlyTyrValHisAlaAspLeuVal 160
 DB GATGTTTGGAGCGCTTGTCTTCTTCATCATGAGGCTATGTCATGCGGACCTCAA 638
 QY 161 ProArgAsnIleLeuTrpSerAlaGluLeuGlyCysPheValLeuIleAspPheGlyLeu 180
 DB CCACGTAACATATTTGTGGAGTGCAGAGATGATTTTAACTATTTGACTTGGACTT 698
 QY 181 SerPheValGluGlyAsnGlnAspValIleValTyrIleGlnThrAspGlyTyrArgAlaPro 200
 DB ACCTTCAAAGAGGACATCAGATGTAAGTATATTCAGACAGAGGCTATGGCTCCA 758
 QY 201 GlnAlaGluLeuGlnAsnGlyAsnGlyLeuAlaGlnGlyLeuGlnSerAspThrGlyCysThr 220
 DB GAAGCAGAAATTCAAATGTGTGGCCAGGCTGGCCGACAGATGATACAGAAATGTACC 818
 QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetCys 240
 DB TCAGCTGTGATCTGTGGAGCTGAGAAATCATTTTACATGGAATGTTCTCAGAAATGAA 878
 QY 241 LeuValHisIleThrValArgSerGlnGluTyrPheValAsnSerSerAlaIleIleAspHis 260
 DB CTGAACATACATCAGATCTCAGAAATGGAAGCAAGTTCGTACTATTTGATCAC 938
 QY 261 IlePheAlaSerValAlaValAlaValAlaIleProAlaTyrHisIleLeuArgAspLeu 280
 DB ATATTGCGAGTAAAGCAGTGGTGAATCCGCAATTCACAGCTTACACTTAAGAACCTT 998
 QY 281 IleValSerMetLeuHisAspAspProSerArgGlyIleProAlaGluMetAlaLeuCys 300
 DB ATCAAAAGCATGCTCATGATGATCCAGCAAAATTCCTGCAAAATGCAATTTGTC 1058
 QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
 DB AGCCCATCTTTAGCATTCCTTTTCCCTCATATTTGAAGATCTGTGATGCTTCCACT 1118
 QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGlnGluGlyTyr 340
 DB CCAGTGTCAAGACTGCTGAATGTCTGATGATGATTTCTTGAAGAAAGAGAGAAATAT 1178
 QY 341 GluAspValValGluAspValIleGluGlyCysGlnIleTyrGlyProValValSerLeu 360
 DB 1178 ----- 1178
 QY 361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
 DB 1179 -----GAGGTCTTTTGTGATGATGCAAAATGCTGG 1208

QY 380 yAepSerLySaLaAGlnlyLeuLeuThRgLyArMePheAspGlyLysPheVala 400
DB 1209 TGAATTCAGAGCTGCGCAGAAATTACTGACTGAGAGATTTGATGGAGATTGTGT 1268
QY 400 lAlaThrPheTyRProLeuSerAlaTyRlySaArgLyTyRLeuTyRGlThRLeuLeu 419
DB 1269 GGCTACATCTACCCCTGAGTGCCTACAGAGGGGATATCTGTATCAAACTTGTCT 1326

RESULT 15
HSM805327 2940 bp mRNA linear PRI 12-JUL-2002
LOCUS Homo sapiens mRNA, cDNA DKFZp6710511 (from clone DKFZp6710511).
DEFINITION AL834136
ACCESSION AL834136
VERSION AL834136.1 GI:21739660
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (base 1 to 2940)
AUTHORS Koehler, K., Beyer, A., Mewes, H.W., Weil, B. and Wiemann, S.
TITLE Direct Submission
JOURNAL Submitted (09-JUL-2002) 1, D-85764 Neuberg, GERMANY
COMMENT Researched from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charite,
Berlin/Germany) within the CDNA sequencing consortium of the German
Genome Project.
This clone (DKFZp6710511) is available at the RZPD in Berlin.
Please contact the RZPD: Resourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further
information about the clone and the sequencing project is available
at <http://mips.gsf.de/proj/cdna/>.
Location/Qualifiers
1..2940
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp6710511"
/db_xref="taxon:9606"
/map="11"
/clone="DKFZp6710511"
/cissue_type="lymph node"
/clone_id="667 (synonym: hln02). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
2923

polYA_site
ORIGIN

Alignment Scores:
pred. No.: 6 49e-163 Length: 2940
Score: 1893.50 Matches: 389
Percent Similarity: 63.46% Conservative: 0
Best Local Similarity: 63.46% Mismatches: 2
Query Match: 85.83% Indels: 223
DB: 9 Gaps: 1

US-10-798-532-2 (1-419) x HSM805327 (1-2940)

QY 29 LeuGlySerGlySerSerAlaSerValTyRArgValaRgCySAspGlyAsnProGlySer 48
DB 1 CTGGGATAGGGGCTCCCTCGCTCGGTGTATCGGGTTCCGCTGCGGCAACCTGTGCTCG 60
QY 49 ProProGlyValAlaLeuLySgInPheLeuProProGlyThRThRGlYAlaAlaAspAla 68
DB 61 CCCCCCGGCGCCCTCAGCAGTCTTGTCCGCCAGGAACACCGGGGCTGCGGCTCTGCC 120
QY 69 AlaGlyTyRGlYpheaRgLySgInaRgAlaAlaLeuGlnGlnGlnGlnGlnGlnGln 88
DB 121 GCGGAGATAGGTTCCGCAAGAGAGGGCGCGCTGGAACAGTTGACGGGTCCAGAAAC 180
QY 89 lLeValThRLeuTyRGlYValPheThRlLehAPheSerProAsnValProSerArG 108
DB 181 ATCGTACCTTGTATGAGAGTGTATTAACAATCACTTTTCTCCAAATGTCATACCGCTGT 240

QY 109 LeuLeuLeuGlnLeuLeuAspValSerValSerGlnLeuLeuTyRSerSerH 128
DB 241 CTGTGTCTGAACCTCTGGATGTCAAGTGTGTGGAAATGCTTATATTCAGTCAACAG 300
QY 129 GlyCySerMetTrpMetlIeGlnHsCySaLaArGAspValLeuGlnAlaLeuAlaPhe 148
DB 301 GGTGTGTCAATGTGATGATACAGCATTTGTGCCAGAGATGTTTGGAGGCGCTTGTCTT 360
QY 149 LeuHlshIeGlnGlyTyRValHlSalAspLeuLyPProArGAsnllLeuTrpSerAla 168
DB 361 CTTCATCATGAGAGGCTATGTCCATGCGGACCTCAACACCATATATTTGGAGTGA 420
QY 169 GluAsnGlnCySPhelyLeuLeuLeaPheGlyLeuSerPheLySgInGln---- 187
DB 421 GAGAAATGAATGTTTAACTCATGACTTGTGACTTGAAGAGGCAATCA-GGT 479
QY 187 ----- 187
DB 480 AAGAAATACCTTTCTTCTCTGCACTTAAATGATGCTTGAAGCCAAATATT 539
QY 187 ----- 187
DB 540 TTAATGACAGGATAGTTAGTAGAGGCTTCAAGTCACTTTAACTTAATTACTGAA 599
QY 187 ----- 187
DB 600 TGCCAGGGATGACGTAATCAAAAAGATTGATGCCACCGTGAATAAAGACATTA 659
QY 187 ----- 187
DB 660 TGTGTGGCTCAGTAGAGAGGAGCAGTGAAGATGAGCTGTGAATTATTAAGTAC 719
QY 187 ----- 187
DB 720 TTAACCTACATGAATAGTTTAAATTTATAGCGTGTGGCTAATGCTGAGTACT 779
QY 187 ----- 187
DB 780 ATTCCTGCTCTCTTCCAGTGTGGGCTGTATCTCGTGAGAAATGATGATGACTTCTG 839
QY 187 ----- 187
DB 840 CCAATTCCTCTTCTCCATGATGTTATCATTTCTTGCAAGTCCCTTATATACAGACAG 899
QY 187 ----- 187
DB 900 CCTAATAGGTAAAGACAAATCCAAAGAGACTTCTTAAGACGCTATATGAAATGTAC 959
QY 187 ----- 187
DB 960 AGCTTGCTCTGAGAACTGCCAGTTAATGAACAAAGATATGTGTAATCAAGACTTA 1019
QY 187 ----- 187
DB 1020 AAATTTTACATTTAGCTTGGCAGTGGGTTACTGCACTTTAGGGAAGTAAATTCAC 1079
QY 187 ----- 187
DB 1080 GTAGAGATGATAGAAAGGAGACAGCTTTTATTGTTGTATATTTACTCATCTTTT 1139
QY 188 ---AspValLyTyRlIeGlnThRAspGlyTyRArgAlaProGlnAlaGlnLeuGlnAn 206
DB 1140 TTAGGATGTAAATATATTCAGACAGCGGATATCGGGCTCCGAAAGCAAAATTTGCAAA 1199
QY 207 CyLeuAlaGlnAlaGlyLeuGlnSerAspThRGlucyThSerAlaValAspLeuTrp 226
DB 1200 TGCTTGCCCAAGGCTGCTGCTGAGACTATACAGATGTACTCAGCTGTATCTGTG 1259
QY 227 SerLeuGlylIeIleuLeuGlnUetPheSerGlyMetLyLeuLySgInThRValArg 246
DB 1260 AGCTTAGAATCATTTTACTGGAATATGTTCTCAGAAATGAACACTGAACATACAGTAC 1319

```
QY 247 SerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHisIlePheAlaSerLysAla 266
    |||
Db 1320 TCTCAGGAATGGAAGGCAACAGTTCGTATTGATTCACATATTGCCAGTAAAGCA 1379
    |||
QY 267 ValValAsnAlaAlaIleProAlaIleThrHisLeuArgAspLeuIleLysSerMetLeuHis 286
    |||
Db 1380 GTGGTAAATGCCGCAATTCCAGCCTTACACCTTAAGAGACCTTATCAAAAGCATGCTTCAT 1439
    |||
QY 287 AspAspProSerArgArgIleProAlaGluMetAlaLeuCysSerProPhePheSerIle 306
    |||
Db 1440 GATGATCCAGACGAAAGAAATTCCTGCTGAATGCAATGTCAGCCCATTTCTTAGCATT 1499
    |||
QY 307 ProPheAlaProHisIleGluAspLeuValMetLeuProThrProValLeuArgLeu 326
    |||
Db 1500 CCTTTGCCCTCATATTGAAGATCTGTCATGCTTCCACTCCAGTGTAAAGACTGCTG 1559
    |||
QY 327 AsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyrGluAspValValGluAsp 346
    |||
Db 1560 AATGTCTGATGATGATTATCTTGAGAAATGAAGAGAAATATGAAGATGTTGTAGAAGAT 1619
    |||
QY 347 ValIleGluGluCysGlnLysTyrGlyProValValSerLeuValProLysGluAsn 366
    |||
Db 1620 GTAAAAGAGAGAGTCTCAAAAATATGACACAGTGTATCTTACTTGTCCAAAGGAAAAAT 1679
    |||
QY 367 ProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGlyAspSerLysAlaAlaGln 386
    |||
Db 1680 CCTGGCAGAGAGACAGTCTTGTGTGAGTATGCAAAATGCTGTGATTCCAAAGCTGGCAG 1739
    |||
QY 387 LysLeuLeuThrGlyArgMetPheAspGlyLysPheValValAlaThrPheTyrProLeu 406
    |||
Db 1740 AAATTACTGACTGAAAGAGATGTTGATGGAGAGTTGTTGTGGCTACATTCTACCCGCTG 1799
    |||
QY 407 SerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeu 419
    |||
Db 1800 AGTGCTACAGAGGGGATATCTGTATCAAACTTGCTT 1838
    |||
```

Search completed: November 29, 2004, 18:10:34

Job time : 4620 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 29, 2004, 11:25:02 ; Search time 502 Seconds

(Without alignments)
4381.489 Million cell updates/sec

Title: US-10-798-532-2

Perfect score: 2206
Sequence: 1 MAGSCGAMAEPRPRLIARG.....VATFYPLSAKXGYLQTL 419

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 413486 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame_p2n.model -DEV=xlp
-Q=/cgm2.1/USPTO.spool.p/US10798532/runat.29112004.112452.15882/app.query.faeta_1.1166
-DB=N Geneseq.23Sep04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPT=0
-LOOPT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=humand0.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10798532.0CGN.1.1.886.0runat.29112004.112452.15882 -NCPU=6 -ICPU=3
-NO MMAP -IARGEOQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq.23Sep04:*
1: geneseqn1980s:*
2: geneseqn1980s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2206	100.0	1260	3	AAZ51355 Human wil
2	2203	99.9	1260	3	AAZ51356 Human tra
3	2198	99.6	1260	4	AAI66829 Human pro
4	2198	99.6	2008	6	ABQ9383 Human cod
5	2198	99.6	2622	4	AAI66828 Human pro
6	2198	99.6	4065	4	AAH72714 Human cer

7	2198	99.6	4065	5	ADL45238 Human ova
8	2191	99.3	1260	4	AA506734 Polynucle
9	2191	99.3	1824	4	AAI18818 Human kin
10	2186	99.1	3244	10	ADBS2857 Primary r
11	2045.5	92.7	3998	12	ADQ24401 Human sof
12	789	35.8	529	4	AAH70764 Human cer
13	766.5	34.7	722	3	AAH02535 Human col
14	729	33.0	470	4	AAH71966 Human cer
15	636	28.8	461	4	AAH69101 Human cer
16	559.5	25.4	411	6	ABQ99225 Human ORF
17	361	16.4	231	3	AAH01259 Human col
18	311.5	14.1	419	5	ADL73740 Human ova
19	311.5	14.1	419	5	ADL67345 Human ova
20	287	13.0	546	6	ABQ53127 Oligonuc
21	287	13.0	546	6	ABQ53126 Oligonuc
22	287	13.0	546	6	ABQ44465 Oligonuc
23	287	13.0	547	6	ABQ44464 Oligonuc
24	288	11.7	546	6	ABQ53129 Oligonuc
25	288	11.7	546	6	ABQ53128 Oligonuc
26	258	11.7	547	6	ABQ44463 Oligonuc
27	258	11.7	547	6	ABQ44462 Oligonuc
28	252	11.4	1065	10	ADD21399 Golden ha
29	250	11.3	442	9	ACH28557 Human adu
30	249.5	11.3	1400	10	ADD21407 Rat cycli
31	248	11.2	1708	10	ADD21395 Mouse cyc
32	243.5	11.0	2472	3	AAc45716 Arabidops
33	243	11.0	1297	10	ADD21393 Human cyc
34	243	11.0	2213	6	ABR84073 Human cDN
35	243	11.0	2213	11	ADD21403 Chinese h
36	243	11.0	2213	11	ADL31711 Human cDN
37	243	11.0	2213	12	ADH17138 Human cdc
38	243	11.0	2328	10	ACC72837 Human can
39	243	11.0	2734	12	ADQ25034 Human sof
40	240	10.9	1209	10	ADD21401 Human sof
41	239.5	10.9	1335	2	AAV71073 CDK2-gree
42	239	10.8	1297	6	ABQ60787 Human 10C
43	239	10.8	1297	10	ADD21391 Human cyc
44	238.5	10.8	1550	12	ADL45326 Rice 180P
45	237.5	10.8	1344	10	ADD21409 Rat cycli

ALIGNMENTS

RESULT 1	AAZ51355	AAZ51355 standard; DNA, 1260 BP.
ID	AAZ51355	standard; DNA, 1260 BP.
XX	AAZ51355;	
AC	AAZ51355;	
XX	06-JUN-2000	(first entry)
DT	06-JUN-2000	(first entry)
XX	Human wild type serine/threonine kinase KIS (hKIS) gene.	
DE	Human wild type serine/threonine kinase KIS (hKIS) gene.	
XX	KIS; human; hKIS; serine/threonine kinase; cell proliferation; G1 phase;	
KW	dominant gene; cyclin-dependent kinase inhibitor; CKI; p27; modulator;	
KW	treatment; cell proliferative disease; vascular disorder; gene therapy;	
KW	atherosclerosis; restenosis; de.	
XX	Homo sapiens.	
OS	Homo sapiens.	
XX	Key	Location/Qualifiers
FX	Key	1.1260
FT	CDS	/tag= a
FT	CDS	/product= "Human wild type KIS protein"
FT	CDS	/transl_except= (pos:556..561, Asn)
PN	MO20001165-A1.	
XX	02-MAR-2000.	
PD	02-MAR-2000.	
XX	20-AUG-1999;	99WO-US018903.
PF	20-AUG-1999;	99WO-US018903.
XX	21-AUG-1998;	98US-0097710P.

XX (NABE/) NABEL G J.
 PA (NABE/) NABEL E G.
 XX Nabel GJ, Nabel EG;
 PI WPI; 2000-237648/20.
 DR P-PSDB; AAY70305.
 XX Novel serine/threonine kinase hKIS polynucleotides and polypeptides used
 PT for inhibiting the cyclin kinase inhibitor p27, and so alter cell
 PT proliferation.
 XX
 PS Claim 1; Page 56; 709p; English.
 CC The present DNA sequence encodes a wild type human KIS (hKIS), isolated
 CC from a human B-cell library, using a yeast two hybrid screening system.
 CC hKIS is a serine/threonine kinase, that acts as an inhibitory kinase of
 CC cyclin-dependent kinase inhibitor (CKI), p27. hKIS is a dominant gene
 CC that controls cell proliferation and is localised predominantly in the
 CC nucleus. It binds to CKI p27 and inhibits its ability to arrest cells in
 CC G1 phase. It has 99% homology to rat serine/threonine protein kinase KIS.
 CC The hKIS sequences are used to modulate cell proliferation and treat cell
 CC proliferative and vascular diseases. The polynucleotide sequence may be
 CC used in gene therapy to treat vascular disorders such as restenosis or
 CC atherosclerosis
 CC
 SQ Sequence 1260 BP; 311 A; 280 C; 335 G; 334 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.75e-209 Length: 1260
 Score: 2206.00 Matches: 419
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0

US-10-798-532-2 (1-419) x AAZ51355 (1-1260)

QY 1 MetAlAGlSerGlyCysAlaATpGlyAlaGluProProaRphLeuGluAlaPheGly 20
 Db 1 ATGGCGGATCCGGCTGCGCTGGGGGCGGAGCCGCGCTTTCTGGAGGCTTCGGG 60
 QY 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
 Db 61 CGGCTGTGGGAGGTAACAGACCGTCTGGGTAGCGGCTCCGCTCGGAGTTCGGGTT 120
 QY 41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60
 Db 121 CGCTGTGGGCGCAACCTGTGCTGCCCCCGGGCCCTCAAGCAGTTCTTGGCCGAGA 180
 QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgPheGluArgAlaAlaLeu 80
 Db 181 ACCACCGGGGCTCGGCTCTGCGCGCGAGTATGTTCCGCAAGAGAGGGCGGCGCTG 240
 QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
 Db 241 GAACAGTTGAGGTCACAGAAACATCGTACCTTTGATGAGTGTTCACATCCACTTT 300
 QY 101 SerProAsnValProSerArgCysLeuLeuLeuGlnLeuLeuAspValSerValSerGlu 120
 Db 301 TCTCCAAAGTGTGCATCACCGCTGTGTTGCTTGAACCTCTGATGTCTAGTGTTCGGA 360
 QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
 Db 361 TTGCTCTTATATTCAGTCCAGCGGGTGTTCATGTGATGATACAGATTCGCCCGCA 420
 QY 141 AspValLeuGlnAlaLeuAlaPheLeuHisHisGlnGlyTyrValHisAlaAspLeuLys 160
 Db 421 GATGTTTGGAGGCGCTGTGCTTTCTTCATCAAGAGGCGATATCCATGCGGACCTCAA 480
 QY 161 ProArgAsnIleLeuTrpSerAlaGluArgGlnCysPheLeuLeuLeuAspPheGlyLeu 180

Db 481 CCACGTAAATATGTGAGTGCAGAGATGATGTTTAACTCATTCGACTTGGACTT 540
 QY 181 SerPheLeuGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
 Db 541 AGCTTCAAAGAGGCAATCAGATGTAAATATTCACAGACAGGATATGGGCTCCA 600
 QY 201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGlyCysThr 220
 Db 601 GAAGCAGATTCCAAAATGCTTGGCCCGGCTGCGCTGACAGATGATCAGAAATGAC 660
 QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
 Db 661 TCAGCTGTGATCTGTGAGCTCAGAAATCATTTTACGAAATGTCTCAGAAATGAAA 720
 QY 241 LeuLysHisThrValArgSerGlnGluTyrLysAlaAsnSerSerAlaIleIleAspHis 260
 Db 721 CTGAAACATACATCAGATCTCAGAAATGAAAGGCAAGTTCGCTATTTATATCAC 780
 QY 261 IlePheAlaSerLysAlaValAlaValAlaIleProAlaTyrHisLeuArgAspLeu 280
 Db 781 ATATTGGCAGTAAAGCAGTGTGATGCCGCAATTCACGCTTACCTAAGACACTT 840
 QY 281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
 Db 841 ATCAAAAGCATCTTCATGATGATCCAGACAGAAATTCCTGTGAATGSCATTTGTC 900
 QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
 Db 901 AGCCCATCTTTAGCATTCCTTTTCCCTCATATTAAGATCTGTATGCTTCCACT 960
 QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspArgTyrLeuGlyAsnGluGluTyr 340
 Db 961 CCAGTGTCAAGACTGCTGATGTCTGATGTATGTTTCTTGGAAATGAAGAGAAATAT 1020
 QY 341 GluAspValValGluAspValLysGluGlyCysGlnLysTyrGlyProValValSerLeu 360
 Db 1021 GAAGATGTGTGAAGATGTAAAGAGAGTGTCAAAATATGACACGTGTATCTCTA 1080
 QY 361 LeuValIleProLysGluAsnProGlyArgGlyGlnValIleValGluTyrAlaAsnAlaGly 380
 Db 1081 CTGTGTTCCAAAGAAATCTCTGCGAGAGCAAGCTTTGTGATATGCAAAATCTGCT 1140
 QY 381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
 Db 1141 GATTCCAAAGCTGCCAGAAATTAATGACTGGAAGAGATTTGATGAGGGAAGTTTGTG 1200
 QY 401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
 Db 1201 GCTACATTTACCCGCTGAGTCTTACAGAGGGGATATCTGTATCAAACTTGTCT 1257

RESULT 2
 AAZ51356
 ID AAZ51356 standard; DNA; 1260 BP.
 AC AAZ51356;
 XX
 DT 06-JUN-2000 (first entry)
 XX
 DE Human transdominant mutant serine/threonine kinase KIS (hKIS) gene.
 XX
 KW KIS; human; hKIS; serine/threonine kinase; cell proliferation; G1 phase;
 KW transdominant gene; mutant; cyclin-dependent kinase inhibitor; CKI; p27;
 KW modulator; treatment; cell proliferative disease; vascular disorder;
 KW gene therapy; atherosclerosis; restenosis; de.
 XX
 OS Homo sapiens.
 XX
 FH Key location/Qualifiers
 FT CDS 1..1260
 FT /tag= a
 FT /product= "Human transdominant mutant KIS protein"
 FT /transl_except= (pos:556..561, Asn)
 FT misc_feature 161

```

FT      /tag= b
FT      /note= "A to G substitution resulting in a Lys to Arg
FT      change in the amino acid sequence"
XX
XX      WO200011165-A1.
XX
XX      02-MAR-2000.
XX
XX      20-AUG-1999; 99WO-US018903.
XX
XX      21-AUG-1998; 98US-0097710P.
XX
XX      (NABE/) NABEL G J.
XX      (NABE/) NABEL E G.
XX
XX      Nabel GJ, Nabel EG;
XX
XX      MPI; 2000-237648/20.
XX      P-PSDB; AAY70306.
XX
XX      Novel serine/threonine kinase hKIS polynucleotides and polypeptides used
XX      for inhibiting the cyclin kinase inhibitor p27, and so alter cell
XX      proliferation.
XX
XX      Claim 7; Page 59; 70pp; English.
XX
XX      The present DNA sequence encodes a transdominant mutant human KIS (hKIS),
XX      constructed by site directed mutagenesis. A single nucleotide
XX      substitution (A to G) results in a lysine to arginine change in the
XX      protein sequence. hKIS is a serine/threonine kinase, that acts as an
XX      inhibitory kinase of cyclin-dependent kinase inhibitor (CKI), p27. hKIS
XX      controls cell proliferation and is localized predominantly in the
XX      nucleus. It binds to CKI p27 and inhibits its ability to arrest cells in
XX      G1 phase. The hKIS sequences are used to modulate cell proliferation and
XX      treat cell proliferative and vascular diseases. The polynucleotide
XX      sequence may be used in gene therapy to treat vascular disorders such as
XX      restenosis or atherosclerosis
XX
XX      Sequence 1260 BP; 310 A; 280 C; 336 G; 334 T; 0 U; 0 Other;
XX
XX      Alignment Scores:
XX      Pred. No.: 9,44e-209 Length: 1260
XX      Score: 2203.00 Matches: 418
XX      Percent Similarity: 100.00% Conservative: 1
XX      Best Local Similarity: 99.76% Mismatches: 0
XX      Query Match: 99.86% Indels: 0
XX      DB: 3 Gaps: 0
XX
XX      US-10-798-532-2 (1-419) x AA251356 (1-1260)
XX
QY      1 MetAlaGlySerGlyCybaIaTTPGlyAlaGluProProArghPheLeuGluAlaPheGly 20
DB      1 ATGGCGGGAATCCGCGCTGCGCGCGCGCGAGCGCGCGCTTTCTGGAGGCGCTTCGCGG 60
QY      21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
DB      61 CGCGTGTGGAGGTACAGACCGGTCTGGAGGAGCGCTTCGCGCTCGATCGGATCGGGT 120
QY      41 ArgCySeGlyAsnProGlySerProGlyAlaLeuLysGlnPheLeuProProGly 60
DB      121 CGCTGGCGGGCAACCTGGCTGCGCCCGCGGCGCTTCAGGCGAGTCTTGGCGCGAGGA 180
QY      61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluAlaAlaLeu 80
DB      181 ACCACCGGGGCTCGCGCTTCGCGCGCGAGATGTTTCCGCAAGAAGAGGGCGCGCTG 240
QY      81 GluGluLeuGlnGlnHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
DB      241 GAACAGTTGACAGGCTCAGAAACATGTAATTGTATGAGAGTCTTACAAATCCACTTT 300
QY      101 SerProAsnValProSerArgCySeLeuLeuGluLeuLeuAspValSerValSerGlu 120
DB      301 TTCCAATGTGTCATCAGCGCTGTGTCTTGAACCTCCGAGATGTCAGTGTTCGGA 360

```

```

QY      121 LeuLeuLeuTyrSerSerHisGlnGlyCySeSerMetTrpMetIleGlnHisCysAlaArg 140
DB      361 TTGCTCTTATATTCCAGTACAGAGGTTTCTCAATGATGATACAGATTCGCCGCCA 420
QY      141 AspValLeuGluValaLeuAlaPheLeuHisHisGlnGlyTyrValHisAlaAspLeuLys 160
DB      421 GATGTTTGGAGGCCCTTCTTTCTTCATCATAGAGGCTATGTCCATCGGACCTCAAA 480
QY      161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCyPheArgLeuLeuLeuAspPheGlyLeu 180
DB      481 CCACTTAAATATGTGAGTGCAGAGTGAATGATGTTTAACTCAATTACTTGGACTT 540
QY      181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
DB      541 AGCTTAAAGAAAGGCAATCAGATGTAAAGTATTTACAGACAGACGGATTCGGCTCA 600
QY      201 GluAlaGluLeuGlnAsnCySeLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
DB      601 GAAGCAGAAATTCAGAAATTCCTGGCCAGGCTGGCTCGAGAGTATACAGAAATGTACC 660
QY      221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
DB      661 TCAGCTGTGATCTGTGAGAGCTCAGGAATCATTTTACTGGAATGTTCTCAGGAATGAAA 720
QY      241 LeuLysHisThrValArgSerGlnGluTyrLysAlaAsnSerSerAlaIleIleAspHis 260
DB      721 CTGAACCAATACAGTCAATCTCGAATGAAAGCAAGATTCCTGCTATTTATGATCAC 780
QY      261 IlePheAlaSerLysAlaValAlaAsnAlaIleProAlaTyrHisIleLeuArgAspLeu 280
DB      781 ATATTTGCCAGTAAAGCAATGAGTGAATGCCGCAATTCAGCTTACCTAAGAGACTT 840
QY      281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
DB      841 ATCAAAAGCATGCTTCATGATGATCCAAACAGAAATCTGCTGAATGCAATGTGTGC 900
QY      301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
DB      901 AGCCCAATCTTAAAGATCTCTTTGCCCCCTCATATTGAAGATCTGTATGCTCCACT 960
QY      321 ProValLeuArgLeuLeuAsnValLeuAspAspTyrLeuGlyAsnGluGlnGlyTyr 340
DB      961 CCAAGTCTAAGACTGCTGATGATGCTGATGATGATTAATCTTGGGAATGAAGGAATAT 1020
QY      341 GluAspValValGluAspValLysGlnGlyCysGlnLysTyrGlyProValValSerLeu 360
DB      1021 GAAGATGTTGTAAAGATGTAAGAGAGAGTGCACAAAATATGACCACTGTATCTTA 1080
QY      361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
DB      1081 CTGTGTCAAAGAAATCTCGCAGAGGACAAAGTCTTGTGAGATCAAAATGCTGGAT 1140
QY      381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
DB      1141 GATTCCAAAGCTCGGAGAAATTAATCACTGAAGAGATTTTGAAGGAAAGTTGTGTG 1200
QY      401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
DB      1201 GCTACATCTTACCCGCTGAGTGCCTCAAGAGGGGATATCTGATCAAACTTGGCTT 1257

```

RESULT 3
 AA166829
 ID AA166829 standard; cDNA; 1260 BP.
 AC AA166829;
 XX
 XX 07-JAN-2002 (first entry)
 DE Human protein kinase polypeptide 13887 coding sequence.
 XX Protein kinase; 3714; 16742; 23546; 13887; cancer; bone disorder; human;
 KM cytosolic; antiinflammatory; immunosuppressive; cardiac; hepatocytrophic;
 KM

KX	gene therapy; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	CDS 1..1260
FT	/.*tag= a
FT	/product= "protein kinase 13887"
PN	WO200173050-A2.
XX	
PD	04-OCT-2001.
XX	
PF	23-MAR-2001; 2001WO-US009483.
XX	
PR	24-MAR-2000; 2000US-0191846P.
PA	(MILL-) MILLENNIUM PHARM INC.
PI	Meyers R;
DR	WPI; 2001-611632/70.
P	P-PSDB; AAG65767.
XX	
PS	Claim 1; Fig 10A-C; 16pp; English.
CC	The invention provides novel human protein kinase polypeptides, 3714,
CC	16742, 23546 and 13887 and nucleic acid molecules encoding them. The
CC	protein kinase polypeptides can be expressed by standard recombinant
CC	methodology. 3714, 16742, 23546 or 13887 nucleic acids and polypeptides
CC	are useful for diagnostic and screening methods to identify subjects at
CC	risk of having cancer or cellular proliferation and/or differentiation
CC	disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and
CC	modulators are useful for the treatment of cancer, particularly colon
CC	cancer or cellular proliferation and/or differentiation disorders. Other
CC	disorders associated with 3714, 16742, 23546 or 13887 expression or
CC	activity that can be treated include bone related disorders, inflammatory
CC	disorders, autoimmune diseases, cardiovascular disorders and liver
CC	diseases. The present sequence represents a human protein kinase
CC	polypeptide 13887 coding sequence
XX	
SQ	Sequence 1260 BP; 312 A; 279 C; 334 G; 335 T; 0 U; 0 Other;
Alignment Scores:	
Pred. No.:	2,96e-208 Length: 1260
Score:	2198.00 Matches: 418
Percent Similarity:	99.76% Conservative: 0
Best Local Similarity:	99.76% Mismatches: 1
Query Match:	99.64% Indels: 0
DB:	Gaps: 0
US-10-798-532-2 (1-419) x AA166829 (1-1260)	
OY	1 MetAlaGlySerGIYCyaAlaTrpGIyAlaGluProArgRhpLeuGluAlaPhagIy 20
Dd	1 ATGGCGGGATCCGGCGCTCGGCCGGGGCGGGAGCCCCCGGTTCCTGGAGGCCTTCGGG 60
OY	21 ArgLeuTTPGlnValGlnSerArgLeuGlySerGIYSerSerAlaSerValTYrArgVal 40
Dd	61 CGGCTTGAGCAGATACAGAGCGCTCTGGGTAGCGGCTCTCCGCTCGGTGTGAGTT 120
OY	41 ArgCySvGgIYAAsnProGlySerProGIyAlaLeuLysGlnPheLeuProGIy 60
Dd	121 CGCTCTGGGGCAACCCTGGCTCGCCCGCGGGCCCTCAAGAGATTCTTCCGCCAGA 180
OY	61 ThrThrngIyAlaAlaAlaSerAlaAlaGlyTyrgIyPheArgIyGgIuArgAlaAlaLeu 80
Dd	181 ACCACCGGGGCTGGCGCTCTCTCCGCGGATGATGTTTTCCGAAGAAGAGGGCGGCTG 240

QY	81	GlulgluLeuglNglyNHsarxAsnll eValThreulYrGlyValPheThrllShIsphe	100
Db	241	GAAGATTGGAGGGGTCAAGAAACATGTGACTTTGTRTGAAGTGTTCACATCCACTTT	300
QY	101	SerProAenValProSerArgCyaleuleuleuGluLeuleuAapValSerValSerGlu	120
Db	301	TCTCCAAATGTGCATCAACGCTGTCGTGTCTTGAACCTCGGATGTGCAGTGTTCGGAA	360
QY	121	LeuleuleuYrSerSerHieGlnGlyCyaseSerMetPheCll eGlnHieCyalaArg	140
Db	361	TTGCTCTTATATTCACAGTCAACAGAGGTGTTCATGTGGAGAATACAGCAATTTGCCCCGA	420
QY	141	AapValLeuGluAlaLeuAlaPheLeuHieShiegluGlyTrValHieAlaAapLeuLys	160
Db	421	GATGTTTGGAGGCCCTTGCTCTTCTTCATATAGGGCTATGTGCCATCGGAACTCAAA	480
QY	161	ProArgAenll eLeuTrpSerAlaGluAaNGluCyapheLysleuIleAapPheGlyLeu	180
Db	481	CCAGTACATATTTGTGGAGTGCAGAGAAATGAATGTTTAAACATATGACTTTGGAACTT	540
QY	181	SerPheLysgluGlyAaNGluAapValLysTrll eGlnTrpAapGlyTrValArgAlaPro	200
Db	541	AGCTTCAAAAGAAAGGCATCAGAAATGTAAAGTATTCAGACAGAGGGGTATCGGGCTTCCA	600
QY	201	GluAlaGluLeuGlnAaNGlyLeuAlaGlnAlaGlyLeuGlnSerAapThrGluCyethr	220
Db	601	GAACGCAAAATTGCAAAATTTGCTTGGCCAGGCTGGCTCGCAGAGATGTATACGAATGTACC	660
QY	221	SerAlaValaAapLeuTrpSerLeuGlyTrll eLeuleuGluMetPheSerGlyMetLys	240
Db	661	TCAGCTGTGTATCTGTGGAGCTTGAAGATCATTTTACGTGAAAGTCTTCAGGAATGAAA	720
QY	241	LeuLysHieThrValArgSerGlnGluTrpValIlaaNGSerSerAlaIleIlaaPHis	260
Db	721	CTGAACCACTACAGTCAGATCTCAGAAATGGAAGGCAACACTTCGCTATATTGATCAC	780
QY	261	IlePheAlaSerLysAlaValaValaAsnAlaIaIleProAlaTrpHieSleuArgAapLeu	280
Db	781	ATATTTCGCAATTAAAGCAGTGTGAATGCCCAATTCAGGCTATCACTTAAAGACCTT	840
QY	281	IleLysSerMetLeuHieAaAapPProSerArgTrIleProAlaGluMetAlaLeuCyS	300
Db	841	ATCAAAAGCATGCTTCATGATGATCCAAAGCAGAAATTCCTGCTGAATGGCAATTGGCC	900
QY	301	SerProPhePheSerIleProPheAlaProHieSlleGluAapLeuValMetLeuProthr	320
Db	901	AGCCCATTTCTTACCATTCCTTTTGCCCCCTCATATTGAAGATCTGTCATCTTCCACT	960
QY	321	ProValLeuArgLeuLeuAaNGValLeuAaPAPAPTrpLeuGlyAaNGluGluLysr	340
Db	961	CCAGTGCCTAAACATCGCTGAATGTCTGGATGATATTAATCTTGAGAAATGAAGAAGAAATAT	1020
QY	341	GluAapValValaGluAapValLysGluGluCyseGlnLysTrpGlyProValValaSerleu	360
Db	1021	GAAGATGTTGTAGAGATGTAAAAAGAGAGTGTCAAAAAATTTGAGACCGTGTGTCTCTA	1080
QY	361	LeuValProLysGluAaNGProGlyArgGlyGlnValaPheValaGluTrpAlaAaAlaGly	380
Db	1081	CTTGTTCCAAAGGAAAAATCTTGAGCAAGACAAAGTCTTTGTTGATGTACAAATGCTGGT	1140
QY	381	AapSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAapGlyLysPheValaVal	400
Db	1141	GATTCCAAAGCTGGCGCAAAATTAATCACTGCTGAAGAGATGTTTGATGGGAATTTGTTGTG	1200
QY	401	AlaThrPheTrpProLeuSerAlaTrpLysArgGlyTrpLeuTrpGlnThrLeuLeu	419
Db	1201	GCTACATTTCTACCGCTGAGTGCCTTACAGAGGGGGAATATCTGATCAAAACCTTGCTT	1257
RESULT 4			
ABQ93383			
ID ABQ93383 standard; cDNA; 2008 BP.			
XX			

AC AB09383; (first entry)
 XX 25-FEB-2003 (first entry)
 DT
 XX Human coding sequence SEQ ID 116.
 DE
 XX Human; expressed sequence tag; EST; haematopoietic disorder;
 XX central nervous system disease; viral infection;
 XX peripheral nervous system disease; non-healing wound; infectious disease;
 XX immune deficiency; immune disorder; bacterial infection; allergy; cancer;
 XX fungal infection; autoimmune disorder; coagulation disorder; neutropenic;
 XX antileukemic; antineoplastic; immunosuppressive; neuroprotective;
 XX cytotoxic; haematologic; virucide; antibacterial; fungicide;
 XX immunostimulant; cerebroprotective; gene therapy; gene; ss.
 XX Homo sapiens.
 OS
 XX WO200259260-A2.
 PN
 XX 01-AUG-2002.
 PD
 XX 16-NOV-2001; 2001MO-US042950.
 PF
 XX 17-NOV-2000; 2000US-00714936.
 PR
 XX (HYSE-) HYSEQ INC.
 PA
 XX Tang YF, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QH;
 PI Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;
 PI WPI; 2002-590824/63.
 DR N-PSDB; ABP64974.
 XX
 XX New isolated polynucleotide, useful in research, diagnostic or
 PT therapeutic methods, e.g. preventing or treating disorders involving
 PT aberrant protein expression or biological activity.
 PT
 XX
 PS Claim 1; SEQ ID NO 116; 394pp; English.
 XX
 XX The present invention relates to novel human coding sequences (AB09383-
 CC AB09386) and proteins (ABP6482-ABP65022). The sequences are useful in
 CC therapeutic, diagnostic and research methods. The polynucleotides may be
 CC used in the field of molecular biology as hybridisation probes, primers
 CC for PCR, for chromosome and gene mapping, for the recombinant production
 CC of protein, or in generation of anti-sense DNA or RNA. The
 CC polynucleotides are useful in diagnostics as expressed sequence tags
 CC (ESTs) for identifying expressed genes or for physical mapping of the
 CC human genome. The proteins may be used as molecular weight markers, or as
 CC nutritional sources or supplements. The proteins may be used to maintain
 CC and expand cell population in a totipotent or pluripotent state
 CC useful for re-engineering damaged or diseased tissues, transplantation,
 CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The
 CC polynucleotides and proteins are useful for preventing, treating or
 CC ameliorating disorders involving aberrant protein expression or
 CC biological activity, e.g. haematopoietic disorders, central/peripheral
 CC nervous system diseases, mechanical and traumatic disorders, non-healing
 CC wounds, immune deficiencies and disorders, infectious diseases caused by
 CC viral, bacterial or fungal infection, autoimmune disorders, allergic
 CC reactions and conditions, coagulation disorders, cancer. The
 CC polynucleotide sequences of the invention were assembled from ESTs
 CC isolated mainly by sequencing by hybridisation, and in some cases,
 CC sequences obtained from one or more public databases. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp://ipo.int/pub/publshd_pct_sequences
 CC
 XX
 SQ Sequence 2008 BP; 491 A; 451 C; 488 G; 578 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 5.86e-208 Length: 2008
 Score: 2198.00 Matches: 418
 Percent Similarity: 99.76% Conservative: 0
 Best Local Similarity: 99.76% Mismatches: 1

Query Match: 99.64% Indels: 0
 DB: 6 Gaps: 0
 US-10-798-532-2 (1-419) x AB09383 (1-2008)
 QY 1 MetAAGlySerGlyCysAlaATPglYalagluProArgPheLeuGluAlaPheGly 20
 DB ANGGCGGAGATCCGGCTCGCCCTGGGCGGAGCGCCGCTTTCTGGAGCGCTTCGG 138
 QY 21 ArgLeuTTPGlnValGlnSerArgLeuGlySerSerAlaSerAlaTyrTrgVal 40
 DB CGGCTGTGCAAGTACAGAGCCGTCTGGGTACCGCTCCGCTGGGTATGCGGTT 198
 QY 41 ArgCysCysGlyAspProGlySerProGlyAlaLeuGlyGlnPheLeuProGly 60
 DB CGCTGCTGGCGAACCTTGCTGCCCCCGCGCCCTCAAGCACTTCTTCCGCGAGA 258
 QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgGluGluAlaLeu 80
 DB ACCACGGGGCTGGCGCTCTGCGCGAGTATGTTTCGCAAGAGAGCGCGCGCTG 318
 QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
 DB GAACAGTTCAGAGGTCAAGAAACATCGACCTTGATAGAGTGTTCACATTCACCTT 378
 QY 101 SerProAsnValProSerArgCysLeuLeuGluLeuAspValSerValSerGlu 120
 DB TCTCCAAATGTGCCATCAGCTGTCTGTTCCTTGAATCTCTGATCTCAGTCTTCGAA 438
 QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
 DB TTGCTCTTATATTCAGTCAACAGGGTGTTCATGTGATGATACAGCATTCGCGCCA 498
 QY 141 AspValLeuGlnAlaLeuAlaPheLeuHisIleGlnGlyTyrValHisAlaLeu 160
 DB 499 GATGTTTGAAGCCCTTCTTCTTCATCAAGAGGCTATGTCATCGGACCTCAAA 558
 QY 161 ProArgAsnIleLeuTyrSerAlaGluGlnGlyCysPheLeuSerIleAspPheGlyLeu 180
 DB 559 CACGTAACATTTGTGGAGTGCAGAGATGATGTTTAACTCATTCATTCGAGCTT 618
 QY 181 SerPheLeuGluGlyAsnGlnAspValTyrIleGlnThrAspGlyTyrTrgAlaPro 200
 DB 619 AGTTTCAAGAAAGGCAATCAGATGTAAATATTCAGACAGACGGGTATCGGCTCCA 678
 QY 201 GluAlaGluLeuGlnAsnGlyLeuAlaGlnAlaGlyLeuGlnSerAspThrGlyCysThr 220
 DB 679 GAAGCAGAAATTCGAAATGCTTGGCCGAGCTGGCTCAGAGTATCAGAAATGTACC 738
 QY 221 SerAlaValAspLeuTyrSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
 DB 739 TCAAGCTGTGATCTGGAGCTCAGAAATCATTTTCTGAATGTTCTCAGGAATGAA 798
 QY 241 LeuLysHisThrValArgSerGlnGlyTyrPylsAlaAsnSerSerAlaIleIleAspHis 260
 DB 799 CTGAACATACATGATGATCTCAGAAATGGAAGGCAACAGTTCTGCTATTTATTCAC 858
 QY 261 IlePheAlaSerLysAlaValAlaAlaAlaIleProAlaTyrHisIleLeuArgAlaLeu 280
 DB 859 ATATTTTCCAGTAAGACAGTGGGATGCGCAATTCAGCTTATCACTTAAGAGCTTT 918
 QY 281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
 DB 919 ATCAAAAGCATGCTTATATGATTCAGAGCAAGAAATTCCTGCTAAATGGCATTTGTC 978
 QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
 DB 979 AGCCCATTTCTTATGATCTCTTTTGGCCCTCATATGAAGATCTGGTCACTGCTCCACT 1038
 QY 321 ProValLeuArgLeuLeuValLeuAlaAspAspTyrLeuGlyAsnGlnGluGluTyr 340
 DB 1039 CCAGTCTAAAGTCTGTAATGTGATGATGATATCTTGAAGATGAAGAGGAATAT 1098

QY 341 GIUAAPVAlValGIUAAPVAlValGIUGluCySGluValTyRGIProValValSerIeu 360
DB 1099 GAAGAGTGTGTAGAGATGTAAAGAGAGTGTCAAAAATATGACACAGTGTATCTCTA 1158
QY 361 LEUValProLYSGluSerProGlyArgGlyGlnValPheValGluValTyrAlaAsnAlaGly 380
DB 1159 CTGTGTCCAAGGAAATCTGGCAGAGCAAGTCTTGTATGATGCAAAATGCTGGT 1218
QY 381 AEPSeRValAlaIaGluValLeuThRGIArgMetPheAPGlyLysPheValVal 400
DB 1219 GATTCAGAGCTCGCAGAAATTAAGTACAGAGATGTTTATGGAAGTTTGTGTG 1278
QY 401 AlaThrPheTyRProLeuSerAlaTyRValArgGlyTyRleuTyRglnThleuVal 419
DB 1279 GCTACATCTACCCGCTGAGTGTCTACAGAGGGGATATCTGTCAAACTTGCTT 1335
RESULT 5
AA166828
ID AA166828 standard; cDNA; 2622 BP.
XX
AC AA166828;
XX
DT 07-JAN-2002 (first entry)
XX
DE Human protein kinase polypeptide 13887 encoding cDNA.
XX
XX Protein kinase; 3714; 16742; 23546; 13887; cancer; bone disorder; human;
KM cytostatic; antiinflammatory; immunosuppressive; cardiac; hepatocrotrophic;
KM gene therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 5'UTR 1..266
FT CDS /tag= a
FT /tag= b
FT /product= "protein kinase 13887"
FT /note= "coding sequence specifically claimed"
FT 3'UTR 1530..2622
FT /tag= c
XX
PN WO200173050-A2.
XX
PD 04-OCT-2001.
XX
PF 23-MAR-2001; 2001MO-US009483.
XX
PR 24-MAR-2000; 2000US-0191846P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Meyers R;
XX
DR WPI; 2001-611632/70.
XX
DR P-PSDB; AAG65767.
XX
XX
PT New human protein kinase polypeptides, 3714, 16742, 23546 and 13887,
PT useful in diagnosis of cancer or cellular proliferation or
PT differentiation disorders and to screen for polypeptide modulators useful
PT to treat such conditions.
XX
XX
PS Claim 1; Fig 10A-C; 169pp; English.
XX
XX The invention provides novel human protein kinase polypeptides, 3714,
CC 16742, 23546 and 13887 and nucleic acid molecules encoding them. The
CC protein kinase polypeptides can be expressed by standard recombinant
CC methodology. 3714, 16742, 23546 or 13887 nucleic acids and polypeptides
CC are useful for diagnostic and screening methods to identify subjects (at
CC risk of) having cancer or cellular proliferation and/or differentiation
CC disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and
CC modulators are useful for the treatment of cancer, particularly colon
CC cancer or cellular proliferation and/or differentiation disorders. Other

CC disorders associated with 3714, 16742, 23546 or 13887 expression or
CC activity that can be treated include bone related disorders, inflammatory
CC disorders, autoimmune diseases, cardiovascular disorders and liver
CC diseases. The present sequence represents a human protein kinase
CC polypeptide 13887 encoding cDNA
XX
SQ Sequence 2622 BP; 652 A; 567 C; 627 G; 776 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 8, 67e-208 Length: 2622
Score: 2198.00 Matches: 418
Percent Similarity: 99.76% Conservative: 0
Best Local Similarity: 99.76% Mismatches: 1
Query Match: 99.64% Indels: 0
DB: 4 Gaps: 0
US-10-798-532-2 (1-419) x AA166828 (1-2622)
QY 1 MetAlaGlySerGlyCySeAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
DB 267 ATGGCGGATCCGGCTGCGCTGCGGCGCGGAGCCGCCGCTTCTGAGGCTTCGGG 326
QY 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
DB 327 CGGCTGGAGAGTACAGACCGCTGAGTACGGCTCTCCGCTCGATGTGCGGTT 386
QY 41 ArgCySeGlyAsnProGlySerProProGlyAlaLeuValGlnPheLeuProProGly 60
DB 387 CGGTGTGGGACACCTGTGGCTGCGCCCGCGGCGCTCAACAGATTTGCGCAGGA 446
QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGlyTyRValPheArgGlyGluArgAlaLeu 80
DB 447 ACCACCGGGCTGCGGCTCTGCGCGCGAGTGTGTTCCGAAAGAGGGCGGCGCTG 506
QY 81 GluGlnLeuGlnGlyValIleArgAsnIleValThrLeuTyRGIValPheThrIleHisPhe 100
DB 507 GAACAGTTCAGAGGTACAGAAACATGTAATGTGTATGAGAGTGTTCATTCACACTT 566
QY 101 SerProAsnValProSerArgCySeLeuLeuGluGluLeuAspValSerValSerGlu 120
DB 567 TCTCCAAATGTGCATCAAGCTGTGTTGTTGAACCTCTGAGTGTCACTGTTCCGAA 626
QY 121 LeuLeuLeuTyRSerSerHisGlnGlyCySerMetTrpMetIleGlnHisCyAlaArg 140
DB 627 TTGCTCTTATATTCAGTACACAGGTTGTTCATGTGATGATACAGCATTTGTCCGA 686
QY 141 AspValLeuGluAlaLeuAlaPheLeuHisGlnGlyTyRValHisAlaAsnLeuVal 160
DB 687 GATGTTTGAAGCCCTTCTCTTTCTTCATCATGAGGCTATGTCCATCGAATCTCAAA 746
QY 161 ProArgAsnIleLeuTrpSerAlaGluAsnGlyCyPheValLeuIleAsnPheGlyLeu 180
DB 747 CCACGTATCATATTTGGAGTGCAGAGATGATGATTTTAACTCATTTGAGACTT 806
QY 181 SerPheValGluGlyValAsnGlnAspValValTyRleGlnThrAspGlyTyRArgAlaPro 200
DB 807 AGCTTCAGAGAGGCAATCAGATGTAAATATTCAGACAGACGAGTATCGGCTCCA 866
QY 201 GIUAAGluLeuGluAsnCySeLeuAlaGlnAlaGlyLeuGlnSerAspThrGlyCyThr 220
DB 867 GAAGCAGAAATGCAAAATTTGCTTGGCCAGGCTGGCTCAAGTATACAGAAATTCAC 926
QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
DB 927 TCAGCTGTGATGTGTGAGAGCTCAGGAATCATTTTATGGAATGTTCACAGAAATGAA 986
QY 241 LeuValHisThrValArgSerGlnGluTrpValAsnSerSerAlaIleIleAsnHis 260
DB 987 CTGAACATATCAGTCACTCAGATCAGGAATGAGAGGCAAAAGTTCGCTTTATTTGATCAC 1046
QY 261 IlePheAlaSerValAlaValAlaAsnAlaIleProAlaTyRHisLeuArgAsnLeu 280
DB 1047 ATATTTCAGATTAAGCAAGTGTGTAATGCGGCAATTCAGCTTATCACTTAAGACCTT 1106

QY 281 ILeYSeMeLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
 DB 1107 ATCAAAAGATGCTTCATATGATGATCCAGAGAGAAATTCCTGCTGAATGGCATTGTGC 1166
 QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
 DB 1167 ACCCATTTCTTATAGCATTCCTTTTGGCCCTCATATTGAAGATCTGGTCAATGGCTCCCACT 1226
 QY 321 ProValLeuArgLeuLeuValLeuAspAspAspPheLeuGluYasnGluGluGluYr 340
 DB 1227 CCAAGTCTAAAGCTGCTGATGTGCTGATGATGATTAATCTTGAGATGAAGAAGGATATAT 1286
 QY 341 GluAspValValGluAspValIlysgIugIuCySGluLysTyrgIyProValIAspLeu 360
 DB 1287 GAAGATGTTGTGAGATGTAAGAGAGGTGTCAAAATATGACCAAGTGTATCTTCA 1346
 QY 361 LeuValProIlysgIuAsnProGlyArgGlyGlnValPheValGluYrAlaAsnAlaGly 380
 DB 1347 CTTGTTCCAAAGGAATCTCTGSCAGAGCAAGCTTTTGTGATGATCAAAATGCTGGT 1406
 QY 381 AspSerIlyAlaAlaGlnIlySleuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
 DB 1407 GATTCCAAAGCTGCGCAGAAATTAAGTGAAGGATGTTGATGGAAAGTTGTGTGG 1466
 QY 401 AlaThrPheTyrgProLeuSerAlaTyrgLysArgGlyTyrgLeuTyrgLinhLeu 419
 DB 1467 GCTACATTCACCCGCTGAGTGCCTACAAAGGGAGATCTGTATCAAACTTGTCTT 1523
 RESULT 6
 AAH72714
 ID AAH72714 standard; cDNA; 4065 BP.
 AC AAH72714;
 XX
 DT 19-SEP-2001 (first entry)
 XX
 DE Human cervical cancer marker nucleic acid 3988.
 KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
 OS Homo sapiens.
 PN WO200142467-A2.
 PD 14-JUN-2001.
 XX
 PF 08-DEC-2000; 2000WO-US033312.
 XX
 PR 08-DEC-1999; 99US-0169681P.
 PR 21-DEC-1999; 99US-0171350P.
 PR 14-MAR-2000; 2000US-0189315P.
 PR 12-MAY-2000; 2000US-0203791P.
 PR 09-JUN-2000; 2000US-0210600P.
 PR 21-JUL-2000; 2000US-0220114P.
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PI Schlegel R, Deeds J, Berger A, Zhao X;
 XX
 DR MPI; 2001-375006/39.
 XX
 PT New isolated nucleic acid for diagnosing and treating cervical cancer and
 CC for assessing and detecting compounds for treating the cancer.
 XX
 PS Claim 1; Page 810-811; 1051pp; English.
 XX
 CC The invention relates to novel genes (AAH68727-AAH73383) associated with
 CC cervical cancer with cytostatic activity. The nucleic acids and encoded
 CC polypeptides are useful: to assess if a patient is afflicted with
 CC cervical cancer or has a pre-malignant condition; to monitor the
 CC progression of cervical cancer or a premalignant condition in a patient;
 CC and to select and/or assess the efficacy of a compound or therapy for

CC inhibiting cervical cancer in a patient. The nucleic acids may also be
 CC useful for gene therapy
 XX
 SQ Sequence 4065 BP; 1068 A; 822 C; 905 G; 1265 T; 0 U; 5 Other:
 Alignment Scores:
 Pred. No.: 1,656-207 Length: 4065
 Score: 2198.00 Matches: 418
 Percent Similarity: 99.76% Conservative: 0
 Best Local Similarity: 99.76% Mismatches: 1
 Query Match: 99.64% Indels: 0
 DB: 4 Gaps: 0
 US-10-798-532-2 (1-419) x AAH72714 (1-4065)
 QY 1 MetAlaGlySerGlyCyAlaAlaThrGlyValAlaGluProProArgPheLeuGluAlaPheGly 20
 DB 268 ATGGGGGATCCGGCTGGCGCTGGGGCGGAGCGCGCGCTTTCTGAGGCTTCGGG 327
 QY 21 ArgLeuThrGlnValAlaGlnSerArgLeuGlySerGlySerAlaSerValTyrgVal 40
 DB 328 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCTCCGCTCGGTATGGGT 387
 QY 41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuYasnPheLeuProProGly 60
 DB 388 CGCTGCTGGCGAACCTGTGGCTGGCCCCCGCGCTCAAGACAGTTCTTGGCCGACAGA 447
 QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrgIyPheValGlyGluArgAlaAlaLeu 80
 DB 448 ACCACCGGGGCTGGCGCTCTGCTCCCGCAGATATGTTCCGCAAGAGGGCGGCGCTG 507
 QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrgIyValPheThrIleHisPhe 100
 DB 508 GAACAGTTGCAGGTCACAGAAACATCTGACTTTGTATGAGTGTATTAATCACTTT 567
 QY 101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuAspValSerValSerGlu 120
 DB 568 TCTCCAAATGTCCATCAAGCTGTGTTGCTGAACCTCGATGATGAGTTCGGA 627
 QY 121 LeuLeuLeuTyrgSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
 DB 628 TTGCTCTTATATTCAGTCCAGCAGAGGATGTTCCATGATGATATCAGGATTTGCCCA 687
 QY 141 AspValLeuGluAlaLeuAlaPheLeuHisIleGluGlyTyrgValHisAlaAspLeuYs 160
 DB 688 GATGTTTGGAGCCCTTGCTTTCTTCATGAGGGCTATGTCATGCGGACTCAAA 747
 QY 161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheYsLeuIleAspPheGlyLeu 180
 DB 748 CCAAGTAAATATTTGTGGAGTGCAGAGATGAATGTTTAAATCATTCGACTTGGACTT 807
 QY 181 SerPheYsGluGlyYasnGlnAspValIlyTyrgIleGlnThrAspGlyTyrgAlaPro 200
 DB 808 ACCTTCAAGAGGCAATCAGATGAAGATATATTCAGACAGAGGGGATGCGGCTCCA 867
 QY 201 GluAlaGluLeuGlnAspCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
 DB 868 GAAGCAGAAATGCAAAATGCTTGGCCAGGCTGGCTGCAGAGATGATACAGAAATGATCC 927
 QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetYs 240
 DB 928 TCAAGTGTGATCTGTGGAGCTCAGGAATCATTTTACGGAATGTCTCAGGAATGANA 987
 QY 241 LeuYsHisThrValArgSerGlnGluTrpIlyValAsnSerSerAlaIleIleAspHis 260
 DB 988 CTGAAGCAATACAGTCAATCTCAGAGATGGAAGGCAACAGTTCGTATTTTGTATCAC 1047
 QY 261 IlePheAlaSerIlyAlaValAlaAsnAlaAlaIleProAlaTyrgHisIleuArgAspLeu 280
 DB 1048 ATATTGTCAGTAAGCAATGTGTGAATGCGGAATTCAGCTTATCACTTAAGAGCTTT 1107
 QY 281 ILeYSeMeLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300

Db 1108 ATCAAAAGCATGCTTCATGATGATCCAGAGAAATTCCTGTAATGCAATTGTGC 1167
Qy SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
Cc 1168 AGCCCAATCTTTAGCATTCCTTTTGGCCCTCATATTGAAATCTGGTCATGCTCCACAT 1227
Db 321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGluAsnGluGluTyr 340
Qy 1228 CCAGTCTAAGCATGCTGGAATGCTGAGATGATGATATTCCTTCAAGATGAAGGAAATAT 1287
Db 341 GluAspValValGluAspValValGluGluGluGluGluGluGluGluGluGluGlu 360
Qy 1288 GAAGAGTGTGTAGAGATGTTAAAGAGAGATGTCATAAAATATGACACAGTGTATCTTA 1347
Db 361 LeuValProlysgLysAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
Qy 1348 CTGTGTTCCAAAGGAAATCTCTGACAGAGCAAGTCTTTTGAATATGCAATGCTGGT 1407
Db 381 AspSerIleValAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
Qy 1408 GATTCCAAAGCTCGCAGAAATTAATGACTGGAAGGATGTTGATGGGAAGTTGTGTG 1467
Db 401 AlaThrPheThrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
Qy 1468 GCTACATTCTACCGCTGAGTGCCTACAGAGGGGATATCTGTATCAAACTTGCTT 1524

RESULT 7

ID ADL45238 standard; DNA; 4065 BP.

ADL45238;

20-MAY-2004 (first entry)

Human ovarian cancer DNA marker #19128.

Human; ovarian cancer; de; tumour; cytostatic; DNA marker.

Homo sapiens.

MO200170979-A2.

27-SEP-2001.

21-MAR-2001; 2001WO-US009126.

21-MAR-2000; 2000US-0191031P.

25-MAY-2000; 2000US-0207124P.

15-JUN-2000; 2000US-0211940P.

07-JUL-2000; 2000US-0216820P.

25-JUL-2000; 2000US-0220661P.

21-DEC-2000; 2000US-0257672P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Lee J, Little J;

WPI; 2001-611502/70.

Novel isolated nucleic acid molecules (markers) overexpressed in ovarian cancer cells as compared to their normal non-cancerous ovarian cells are used to characterize stage, grade, histological type of ovarian cancer.

Disclosure; SEQ ID NO 19128; 1066P; English.

The invention relates to nucleic acid markers which are overexpressed in ovarian cancer cells as compared to their expression in normal (i.e. non-cancerous) ovarian cells. The invention also relates to polypeptides encoded by the markers, antibodies that selectively bind to the polypeptides, a method of inhibiting ovarian cancer in a patient at risk of developing ovarian cancer involving inhibiting expression of a gene corresponding to a marker of the invention and a method of treating a patient afflicted with ovarian cancer comprising providing to cells of

the patient an antisense oligonucleotide complementary to a marker of the invention. The markers are useful for assessing if a patient is afflicted with ovarian cancer, which involves comparing the level of expression of a marker in a patient sample and a normal level of expression of the marker in a control non-ovarian cancer sample. A difference between the expression levels indicates ovarian cancer. The level of expression of a marker corresponds to a secreted protein or to a transcribed polynucleotide or its portion. The level of expression of the marker is assessed by detecting the presence in the sample, a protein or protein fragment corresponding to the marker. The presence of protein or protein fragment is detected using an antibody that specifically binds with the protein or protein fragment. Alternatively, the level of expression of the marker is assessed by detecting the presence of a transcribed polynucleotide which anneals with the marker or anneals with a portion of the polynucleotide comprising the marker, under stringent conditions. The marker is also used for monitoring the progression of ovarian cancer in a patient which involves detecting expression of the marker in a patient sample at a first point in time, repeating the method at a subsequent time and comparing the level of expression. The method is carried out using an ovarian tissue sample. A composition comprising a marker, polypeptide or antibody of the invention is used to treat ovarian cancer. This sequence represents a human ovarian cancer DNA marker of the invention.

Sequence 4065 BP; 1068 A; 822 C; 905 G; 1265 T; 0 U; 5 Other;

Alignment Scores:

Pred. No.:	1,65e-207	Length:	4065
Score:	2198.00	Matches:	418
Percent Similarity:	99.764	Conservative:	0
Best Local Similarity:	99.764	Mismatches:	1
Query Match:	99.648	Indels:	0
DB:	5	Gaps:	0

US-10-798-532-2 (1-419) x ADL45238 (1-4065)

Qy 1 MetAlaGlySerGlyCysAlaThrProGlyAlaGluProProArgPheLeuGluAlaPheGly 20
Db 268 ATGGCGGATCCGGCTGCGCGCGGAGCGCCGCTTTCTTGAGGCTTCGGG 327
Qy 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
Db 328 CGGCTGTGCGAGTACAGAGCCGCTCTGGGTACGGCTCTCGGCTCGGTATCGGTT 387
Qy 41 ArgCysGlyAsnProGlySerProGlyAlaLeuLysGlnPheLeuProProGly 60
Db 388 CGGTGTGCGGCAACCTGCTGCGCCCGCGCCCTCAACAGATCTTGGCGCAGGA 447
Qy 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
Db 448 ACCACCGGGGCTGGCGCTCTGCGCGCGAGTATGTTCCGAAAGAGGGGCGCGCTG 507
Qy 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
Db 508 GAACAGTTCAGGGTCAAGAAACATCGATCTTGTATGAGAGTGTATCAATCCACTTT 567
Qy 101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuAspValSerValSerGlu 120
Db 568 TCTCCAAATGTGCATCAAGCTGTCTGTGTAACCTCGATGATCACTGTTCCGAA 627
Qy 121 LeuLeuLeuTyrSerSerIleGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
Db 628 TTGCTCTTATATTCAGTACACAGGGTGTCTCAATGATATATACAGATTTGCCCA 687
Qy 141 AspValLeuGluAlaLeuAlaPheLeuHisGluGlyTyrValHisAlaAspLeuLys 160
Db 688 GATGTTTGAAGGCTTCTTTCTTCAATCAAGAGGATATGTCATCGGACCTCAAA 747
Qy 161 ProArgAsnIleLeuTrpSerArgLeuGluGluGluCysPheLysLeuIleAspPheGlyLeu 180
Db 748 CCACGTAATATTTGAGTGCAGAGATGATGATGTTTAACTCAATGACTTTGGACTT 807
Qy 181 SerPheLysGluGluGluAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200

```

DB 808 AGCTTAAAGAGGACATCAGATGTAATATTTTCAAGCAGCGGATATGGGCTCCA 867
QY 201 GUAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 220
DB 868 GAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 927
QY 221 SGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
DB 928 TCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 987
QY 241 LGLYAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 260
DB 988 CTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1047
QY 261 LGLYAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 280
DB 1048 ATATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1107
QY 281 LGLYAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
DB 1108 ATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1167
QY 301 SGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 320
DB 1168 AGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1227
QY 321 PGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 340
DB 1228 CAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1287
QY 341 GUAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
DB 1288 GAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1347
QY 361 LGLYAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 380
DB 1348 GTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1407
QY 381 APTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 400
DB 1408 GATTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1467
QY 401 AATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 419
DB 1468 GCTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1524

```

RESULT 8
AAS06734
ID AAS06734 standard; cDNA; 1260 BP.
XX AAS06734;
AC 12-SEP-2001 (first entry)
DT 12-SEP-2001 (first entry)
XX
DB Polynucleotide sequence encoding human protein kinase #34.
XX
XX Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
XX metabolic disorder; immune related disease; neurological disorder;
XX neurodegenerative disorder; inflammatory disorder; infectious disease;
XX reproductive disorder; gene therapy; ss.
OS Homo sapiens.
XX
XX WO200138503-A2.
XX
XX 31-MAY-2001.
XX
XX 22-NOV-2000; 2000MO-US032085.
XX
XX 24-NOV-1999; 99US-0167482P.
XX
XX (SUB-) SUGEN INC.

```

XX Plozman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
PI Flanagan P, Clary D;
XX WPI; 2001-343950/36.
DR P-PSDB; AAU03534.
XX
PT Nucleic acids encoding human kinase polypeptides, useful for preventing
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT neuronal-associated diseases, and microbial infections.
XX
PS Example 1; Fig 1; 433bp; English.
XX
CC AAS06701-AAS06757 encode for novel human protein kinases #1-57. The novel
CC protein kinases have been identified as members of the tyrosine or
CC serine/threonine kinase (PTK and STK) families. The polynucleotides
CC encoding protein kinases and the polypeptides may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate kinase expression. For example, they may be used to treat
CC cancers (especially cancers of haematopoietic origin), cardiovascular
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC immune related diseases (e.g. rheumatoid arthritis), neurological
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC Additionally, polynucleotides encoding protein kinases may be used for
CC gene therapy and as DNA probes in diagnostic assays. The protein kinase
CC polypeptides may be used as antigens in the production of antibodies
CC against the protein kinases and in assays to identify modulators of
CC protein kinase expression and activity
XX
SQ Sequence 1260 BP; 311 A; 279 C; 335 G; 335 T; 0 U; 0 Other;

```

Alignment Scores:
Pred. No.: 1,476-207 Length: 1260
Score: 2191.00 Matches: 417
Percent Similarity: 99.52% Conservative: 0
Best Local Similarity: 99.52% Mismatches: 2
Query Match: 99.32% Indels: 0
DB: 4 Gaps: 0

US-10-798-532-2 (1-419) x AAS06734 (1-1260)

```

QY 1 MetAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 20
DB 1 ATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
QY 21 ArgLeuTPGlnValGlnSerArgLeuGlnSerArgLeuGlnSerArgLeuGlnSerArgVal 40
DB 61 CGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 41 ArgCysCysGlyAsnProGlySerProGlyAlaLeuGlnPheLeuProGly 60
DB 121 CGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 61 ThrThrGlnValAlaAlaSerAlaAlaGlnValGlnPheArgGlnValAlaLeu 80
DB 181 ACCACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 81 GlnGlnLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 100
DB 241 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 101 SerProAsnValProSerArgCysLeuLeuGlnLeuLeuAspValSerValSerGln 120
DB 301 TCTCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 121 LeuLeuLeuTyrSerSerHisGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 140
DB 361 TTGCTTTATATTCAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 141 AspValLeuGlnValAlaLeuAlaPheLeuHisGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 160

```

Db 421 GATGTTTGGAGGCCCTGCTTTCTTCATCATAGAGGCTAGTCCATGCGGACCTCAAA 480
 QY 161 ProArgAsnIleLeuTrpSerAlaGluAenGluCysPheLeuIleAspPheGlyLeu 180
 Db 481 CCACGTAACATATGTCGAGCTGAGAGATGATGTTTAACTCATTTGACTTGGACTT 540
 QY 181 SerPheLeuGluGlyAenGluAenValIleValIleGlnThrAspGlyTyrArgAlaPro 200
 Db 541 AGCTTCAAAGAGGCAATCAGATGTAATGTAATATTCAGACAGACGGGATCCGGCTCCA 600
 QY 201 GUAAGluGluGlnAenCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGlyCysThr 220
 Db 601 GAAGCGAATGTAATAATGCTTGCGCCAGAGCTGCGCTGAGTAATACAGATATACC 660
 QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetCys 240
 Db 661 TCAGCTGTTGATGTGGAGCCCTAGGAATCATTTTCTGGAATGTTCTCAGGAATGAAA 720
 QY 241 LeuValHisThrValArgSerGlnGluTrpValAsnSerSerAlaIleIleAspHis 260
 Db 721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGCAACAGTTCCTCATTTATGATCAC 780
 QY 261 IlePheAlaSerIleValAlaValAlaIleProAlaTyrHisLeuArgAspLeu 280
 Db 781 ATATTGCGCAGTAAGCAGTGTGAATGCCGCAATTCAGCTCATACCTAGAGACCTT 840
 QY 281 IleLeuSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
 Db 841 ATCAAAAGCATGCTTCATGATGATCCAGAGCAAGAAATTCCTGTAATGGCATTTGTCG 900
 QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
 Db 901 AGCCCATCTCTTAAGCATCTCTTGGCCCTCATATGGAAGATCTGCTCATGCTCCACT 960
 QY 321 ProValLeuArgLeuLeuAenValIleAspAspAspTyrLeuGluAenGluGluTyr 340
 Db 961 CAGTCTTAAGATGCTGAATGTGCTGAGATGATGATTAATCTTGAGATGAGAGAAAT 1020
 QY 341 GluAspValAlaGluAspValIleGluGluGluCysGlnIleValIleProValAlaSerLeu 360
 Db 1021 GAAGATGTTGTGAAGATGTAAGAGAGAGCTGCAGAAATATGACACAGTGGATCTCA 1080
 QY 361 LeuValProValGluAenProGlyArgGlyGlnValIleValIleGluTyrAlaAsnAlaGly 380
 Db 1081 CTGTGTCCAAAGGGAATCTCGCAGAGGCAAGTCTTGTGATGATGCAATGCTGGAT 1140
 QY 381 AspSerValAlaIleGlnIleLeuLeuThrGlyArgMetPheAspGlyAspPheValAla 400
 Db 1141 GATTCGAAAGCTGCGAGAAATTAAGTGAAGAGATGTTGATGAGAAATTTGTTGAG 1200
 QY 401 AlaThrPheTyrProLeuSerAlaTyrIleArgGlyTyrLeuValIleGlnThrLeuLeu 419
 Db 1201 GGTACATTTACCGGTGAGTGTCTTCAAGAGGGGATATCTGATCAAACTTGCTT 1257
 RESULT 9
 AAD18818
 ID AAD18818 standard; cDNA; 1824 BP.
 AC AAD18818;
 XX 18-DEC-2001 (first entry)
 DT 18-DEC-2001 (first entry)
 XX Human kinase (PKIN) -3 cDNA.
 DE Human kinase; PKIN; gene therapy; adenocarcinoma; immune disorder; gout;
 XX cancer; allergy; sarcoma; leukemia; acquired immune deficiency syndrome;
 KW AIDS; Addison's disease; microbial infection; inflammation; osteoporosis;
 KW atherosclerosis; cardiovascular disease; myocardial infarction; anaemia;
 KW myaesthesia gravis; cirrhosis; cataract; growth and development disorder;
 KW seizure disorder; pulmonary embolism; Gaucher's disease; lipid disorder;
 KW lipid storage disease; Pick's disease; Tay-Sachs disease; renal disease;
 KW asthma; obesity; restorative therapy; cytostatic; immunomodulatory;
 KW antimicrobial; cardiovascular; antiinflammatory; vaccine; ss.

XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH CDS 169..1428
 FT /tag= a
 FT /product= "Human PKIN-3 protein"
 XX
 PN W0200181555-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US012992.
 XX
 PR 20-APR-2000; 2000US-0199021P.
 PR 28-APR-2000; 2000US-0200226P.
 PR 05-MAY-2000; 2000US-0202339P.
 PR 11-MAY-2000; 2000US-0203505P.
 PR 18-MAY-2000; 2000US-0205564P.
 PR 26-MAY-2000; 2000US-0207739P.
 PR 01-JUN-2000; 2000US-0208795P.
 XX
 PA (INCYTE) INCYTE GENOMICS INC.
 XX
 PI Yue H, Gandhi AR, Trilivouley CM, Kearney L, Griffin JA, Nguyen DB,
 PI Bandman O, Lu DM, Lal P, Burford N, Khan FA, Walla NK, Yao MG,
 PI Paterson C, Burrill JD, Marcus GA, Zingler KA, Reclon SA, Lu Y,
 PI Policky JL, Thornton M, Tang YT, Hafalia A, Elliott VS, Baughn MR,
 PI Walsh RF, Rankumar J, Borowsky ML, Au-Young J, Hillman JL,
 PI Gururajan R;
 XX
 DR WPI; 2001-611740/70.
 DR P-PSDB; AAE11769.
 XX
 PT Human kinases and nucleic acids, useful for preventing diagnosing and
 PT treating cancers, inflammation and immune disorders.
 XX
 PS Claim 5; Page 152; 166pp; English.
 XX
 CC The present invention relates to human kinases (PKIN) and the nucleic
 CC acids encoding them. PKIN is used as vaccine and in gene therapy. PKIN is
 CC used in the prevention, diagnosis and treatment of diseases cancers,
 CC adenocarcinoma, leukemia, sarcoma, immune disorder, Addison's disease,
 CC acquired immune deficiency syndrome (AIDS), anaemia, asthma, allergies,
 CC gout, microbial infections, cardiovascular disease and/or inflammation,
 CC myaesthesia gravis, atherosclerosis, cirrhosis, osteoporosis, myocardial
 CC infarction, cataract, growth and development disorder, seizure disorder,
 CC pulmonary embolism, Gaucher's disease, lipid disorder, lipid storage
 CC disease, Pick's disease, Tay-Sachs disease, renal disease and obesity.
 CC PKIN may be used to treat disorders associated with decreased PKIN
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of PKIN by expressing inactive proteins or to
 CC supplement the patient's own production of PKIN. PKIN nucleic acids may be
 CC used to produce the PKIN polypeptide, by inserting the nucleic acids into
 CC a host cell and culturing the cell to express the protein. PKIN nucleic
 CC acid and its complementary sequences may also be used as DNA probes in
 CC diagnostic assays to detect and quantitate the presence of similar
 CC nucleic acid sequences in samples and therefore which patients may be in
 CC need of restorative therapy. The present sequence is human PKIN-3 cDNA
 XX
 SQ Sequence 1824 BP; 434 A; 416 C; 472 G; 502 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2,53e-207 Length: 1824
 Score: 2191.00 Matches: 417
 Percent Similarity: 99.52% Conservative: 0
 Best Local Similarity: 99.52% Mismatches: 2
 Query Match: 99.32% Indels: 0
 DB: 4 Gaps: 0
 US-10-798-532-2 (1-419) x AAD18818 (1-1824)
 QY 1 MetaGlySerGlyCysAlaTrpGlyAlaGluProArgPheLeuGluAlaPheGly 20

```

Db      169 ATGGCGGGAATCCGGCTGCGCTGGCGCGGAGACCCCGGCTTTCTGAGGCGCTTCGGG 228
Qy      21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
Db      229 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCTCCGCTCGGTGTTCGGGTT 288
Qy      41 ArgCysCysGlyAsnProGlySerProProGlyValAlaLeuLysGlnLeuProProGly 60
Db      289 CGGTGTCGGCAACCTGTGCTGCCCCCGGCGCCCTCAAGAGATCTTGTGCGCAGGA 348
Qy      61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgPheGluArgAlaAlaLeu 80
Db      349 ACCACCGGGGCTGCGGCTTCTGCGCGGAGTATGTTTCCGCAAGAGAGGGCGGCTG 408
Qy      81 GlnGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
Db      409 GAACAGTTGGCAGGGTCACAGAAACATCGTGACTTTGTATGAGTGCTTTCATCAATCCACTTT 468
Qy      101 SerProAsnValProSerArgCysLeuLeuLeuGlnLeuLeuAspValSerValSerGlu 120
Db      469 TCTCCAAATGTGCATCACGCTGTCTGTGCTTGAACCTCTGATGTCAGTGTTTGGAA 528
Qy      121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
Db      529 TTGCTCTTATATTCAGTCCAGGGTGTTCATGTGATGATGATGATGATGATGATGATGATG 588
Qy      141 AspValLeuGlnAlaLeuAlaPheLeuHisHisGlnGlyTyrValHisAlaAspLeuLys 160
Db      589 GATGTTTGGAGGCGCTTGTCTTCTTCATCATGAGGCTATGTCATGCGAGCTCAAA 648
Qy      161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
Db      649 CCACGTAACATATTTGGAGTGGAGAGAAAGTAATGTTTAACTCATATGACTTGGACTT 708
Qy      181 SerPheLysGlnGlyAsnGlnAspValLysTyrCileGlnThrAspGlyTyrArgAlaPro 200
Db      709 AGCTTCAAGAGAGGCAATCGAGATGTAATGTAATTCAGACAGCGGATTCGGGCTCCA 768
Qy      201 GlnAlaGlnLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspTrpGluCysThr 220
Db      769 GAAGCAGAAATTCGAAATTCCTTGCCCGAGGCTGCGCTGAGAGTATACAGAAATGACC 828
Qy      221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnMetPheSerGlyMetLys 240
Db      829 TAGCGTGTGATCTGGAGACCTCAGAAATATTTACTGAAATGTTCTCAGAAATGAAA 888
Qy      241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
Db      889 CTGAACATACATCAGATCTTCAGAAATGAAAGGCAAAACGTTCTGCTATTATTGATCAC 948
Qy      261 IlePheAlaSerLysAlaValAlaAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
Db      949 ATATTGTCAGTAAGCAGTGAAGTGAATGCGCAATTCAGCTATATCACTTAAGACCTT 1008
Qy      281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGlnMetAlaLeuCys 300
Db      1009 ATCAAAAGCATGCTTCATGATGATCCAAAGCAAGAAATTCCTCTGAAAGGCAATGTGC 1068
Qy      301 SerProPhePheSerIleProPheAlaProHisIleGlnAspLeuValMetLeuProThr 320
Db      1069 AGCCCATTTCTTTAGCATTTCTTTTGGCCCTCATATTGAAAGATCTGTGATCTCCACT 1128
Qy      321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlnGlnGlnGlnGlyTyr 340
Db      1129 CCAAGTCTTAAGACTGCGTGAATGCTGTGATGATGATTAATCTTGAATGAAGAGAAATAT 1188
Qy      341 GlnAspValValGlnAspValLysGlnGluCysGlnLysTyrGlyProValValSerLeu 360
Db      1189 GAAGAGTGTGTAAAGATGTAAGAGAGAGAGTGCAAAATAATGACCAAGTGTATCTCTA 1248
Qy      361 LeuValProLysGlnAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380

```

```

Db      1249 CTTGTTCCAAAGGAATCTGCGACAGACAAGTCTTTGTGATGATCAAACTGGT 1308
Qy      381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
Db      1309 GATTCCAAAGCTGCCACAGAAATTAATCTGACAGAGAGATGTTGATGGAAAGTTTGTG 1368
Qy      401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
Db      1369 GCTACATCTACCGCTGATGCTCTTACAGAGGGAATCTGTATCAAACTTGCCTT 1425

RESULT 10
ADBS2857
ID ADBS2857 standard; DNA; 3244 BP.
AC ADBS2857;
AC
DT
DT
DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3399.
KW toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW toxicity marker; toxicity progression; drug screening;
KW primary rat hepatocyte toxicity modelling; gene; db.
OS Rattus norvegicus.
PN
PN MO2003065993-A2.
PD
PD 14-AUG-2003.
XX
XX
PF 04-FEB-2003; 2003WC-US003482.
XX
XX
PR 04-FEB-2002; 2002US-0353171P.
PR 13-MAR-2002; 2002US-0363534P.
PR 08-APR-2002; 2002US-0370248P.
PR 10-APR-2002; 2002US-0371134P.
PR 10-APR-2002; 2002US-0371135P.
PR 10-APR-2002; 2002US-0371150P.
PR 11-APR-2002; 2002US-0371413P.
PR 19-APR-2002; 2002US-0373601P.
PR 22-APR-2002; 2002US-0374139P.
PR 08-MAY-2002; 2002US-0378370P.
PR 09-MAY-2002; 2002US-0378652P.
PR 09-MAY-2002; 2002US-0378653P.
PR 09-MAY-2002; 2002US-0378665P.
PR 09-JUL-2002; 2002US-0394230P.
PR 09-JUL-2002; 2002US-0394253P.
PR 04-SRP-2002; 2002US-0407688P.
PR 28-JUN-2003; 2003US-0442900P.
XX
XX (GENE-) GENE LOGIC INC.
PA
PA Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
PI Elashoff M;
PI
XX
XX WPI, 2003-731472/69.
XX
XX
XX
XX Claim 44; SEQ ID NO 3399; 874bp; English.
XX
XX The present invention describes a method for determining whether a
XX compound induces a toxic effect on a tissue or cell. The method comprises
XX preparing a gene expression profile of a tissue or cell sample exposed to
XX the compound, and comparing the gene expression profile to a database
XX comprising data or information on the Tox mean and non-Tox mean values.
XX The method is useful for predicting or identifying at least one toxic
XX effect, particularly hepatotoxicity, of a test or unknown compound. The
XX genes listed in the specification are useful as diagnostic or toxicity

```

CC markers for the prediction or identification of the physiological state of tissue or cell sample that has been exposed to a compound, or to identify or predict the toxic effects of a compound or an agent. These CC may also be used as markers for monitoring toxicity progression or for drug screening. The present sequence represents a primary rat hepatocyte CC toxicity modelling related gene sequence from the present invention.

XX Sequence 3244 BP; 800 A; 710 C; 819 G; 915 T; 0 U; 0 Other;

Alignment Scores:

pred. No.:	1,84e-206	Length:	3244
Score:	2186.00	Matches:	415
Percent Similarity:	99.52%	Conservative:	2
Best Local Similarity:	99.05%	Mismatches:	2
Query Match:	99.09%	Indels:	0
DB:	10	Gaps:	0

US-10-798-532-2 (1-419) x ADB52857 (1-3244)

```

QY 1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProPheLeuGluAlaPheGly 20
DB 162 ATGGCCGGGTCGGCTGGCGTGGGGGCGGAGCCGCCGCTTCTGGAGGCTTCGGG 221
QY 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
DB 222 CGGCTGTGGCAGGTACAGAGCCGCTTAGGCGAGCGGCTCTCCGCGCTCGGTGACCGGGTG 281
QY 41 ArgCysGlyValAspProGlySerProGlyValAlaLeuGlyGlnPheLeuProGly 60
DB 282 CGCTGCTGGCGACTCCAGGCTGCCGCCGCCGCCCTCAAGAGTTCCTGCTCCGGGA 341
QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgGlyGluArgAlaLeu 80
DB 342 ACCACCGGGGCTCCGCTGGCTGGCGAGTATGTTTCGAAAGAGAGGCGAGCGCTG 401
QY 81 GluGlnLeuGlnGlnValArgAsnIleValThrLeuTyrGlyValaPheThrIleHisPhe 100
DB 402 GAGCAGTTGGAGGGTACAGAGAACATCGTACTTTATACGAGACTCTTACCATACACTTC 461
QY 101 SerProAsnValProSerArgCysLeuLeuGlnGlnLeuAspValSerValSerGln 120
DB 462 TTCCTCAAAATGTCACACGCTCTGTTGCTTGAACCTCTGATGACGCTTTCGGA 521
QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
DB 522 TTGCTGTATATTCAGTCCAGGCTGCTCATGATGATGATCCAGCAGCTGTGCGA 581
QY 141 AspValLeuGlnAlaLeuAlaPheLeuHisGlnGlyTyrValHisAlaAspLeuVal 160
DB 582 GATGCTCTGGAGGCCCTTGCTTTTCTTCAACATAGAGGCTATGTCATGACAGCTCAAG 641
QY 161 ProArgAsnIleLeuTrpSerArgIleGluAsnGluCysPheLeuLeuAspPheGlyLeu 180
DB 642 CCACGAAACATCTCTGTGAGTGGCGAGAACGATGCTTTAAGCTTTATGACTTGGACATC 701
QY 181 SerPheLeuGlnGlyAsnGlnAspValIleTyrIleGlnThrAspGlyTyrArgAlaPro 200
DB 702 AGCTTCAAGAGGCAATCAGACGTAGATATATTCAGACAGGATATCGAGTCTCT 761
QY 201 GluAlaGlnLeuGlnAsnGlyLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
DB 762 GAAGCAGAACTGAGAAATGCTTGGCCCGGCTGAGAGTATACAGAGTGTACC 821
QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnIleuSerPheSerGlyMetCys 240
DB 822 TACGCTGTGATCTTCGAGCTTGGAAATCATTTTACTGGAATGTCTTCAGAAAGAA 881
QY 241 LeuValHisIleThrValArgSerGlnGluTyrPlyValaAsnSerSerAlaIleIleAspHis 260
DB 882 CTGAACCATPACAGTCAATCTCAGAGTGAAGGCAACAGTTCTCTATATATGATCAT 941
QY 261 IlePheAlaSerIleValAlaValaAlaIleProAlaTyrHisLeuAlaGlyLeu 280

```

```

DB 942 ATAATTGCCAGTAAGACGTGGTGAATGCCGCAATTCAGCTTACCTCAGACACTT 1001
QY 281 IleYsSerMetLeuHisAspAspProSerArgAlaIleProAlaGluMetAlaLeuCys 300
DB 1002 ATCAAAAGCATCTTCATGACACACCAAGCAGAAAGATCCCTGCTGAGATGGCTTGGTC 1061
QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
DB 1062 AGCCCATCTTTAGCATTCCTTTGCCCTCATATTGAAGATCTGTGATGCTTCAACT 1121
QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspArgTyrLeuGlyValAsnGluGluTyr 340
DB 1122 CCAGTGCTCAGACTCTCATATGCTGGATGATGACTATCTTGAATAAGATATAAT 1181
QY 341 GluAspValAlaGluAspValIleGlnGluCysGlnTyrGlyProValIleSerLeu 360
DB 1182 GAAAGATGTTGTGAAGATGTAAAGAGTGTCAGAATATGACCAAGTGGTCTCTG 1241
QY 361 LeuValProValProValAspProGlyValArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
DB 1242 CTGTGTTCCAAAGAAATCTTGGCAGAGGACAAATCTTGTGAGTACGCAAACTGGA 1301
QY 381 AspSerIleValAlaGlnIleLeuLeuThrGlyValArgMetPheAspGlyValPheValVal 400
DB 1302 GATTCAAAGCTGCTCAGAAATGCTGACTGGAGAGATGTTGACGGAAATGTGTGTG 1361
QY 401 AlaThrPheTyrProLeuSerAlaTyrIleYsArgGlyTyrLeuTyrGlnThrLeuLeu 419
DB 1362 GCTACATCTTACCGCTGATGCTCTACAAAGAGGAGATATCTTTATCAAACTTGCTT 1418

```

RESULT 11

ADQ24401

ID ADQ24401 standard; DNA; 3998 BP.

XX

AC ADQ24401;

XX

DT 26-AUG-2004 (first entry)

XX

DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 7221.

XX

KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human; db.

XX

OS Homo sapiens.

XX

PN WO2004048938-A2.

XX

PD 10-JUN-2004.

XX

PF 26-NOV-2003; 2003WO-US038193.

XX

PR 26-NOV-2002; 2002US-0429739P.

XX

PA (PROT-) PROTEIN DESIGN LABS INC.

XX

PI Aziz N, Ginsburg WM, Zlocznik A;

XX

DR WPI; 2004-441208/41.

XX

PT Early detection of soft tissue sarcoma comprises determining expression of a gene in a first soft tissue sample and a normal soft tissue sample and comparing the gene expression, also useful in treating soft tissue sarcoma.

XX

PS Example 2; SEQ ID NO 7221; 210pp; English.

XX

CC The invention relates to a novel method for detecting soft tissue sarcoma which comprises obtaining a first soft tissue sample from an individual, CC and a normal soft tissue sample from the same or different individual, CC determining the expression of a gene in both samples and comparing the CC expression of the gene in both soft tissue samples, where a higher level of protein expression in the first soft tissue sample indicates the presence of soft tissue sarcoma. The method of the invention has

CC cytostatic applications and may be useful for detecting soft tissue
 CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
 CC acid sequences may be useful in diagnostic and screening applications.
 CC The current sequence is that of a human soft tissue sarcoma-upregulated
 CC DNA of the invention. The current sequence is not shown within the
 CC specification per se but was submitted in CD format by the inventor.

XX Sequence 3998 BP; 1066 A; 777 C; 885 G; 1255 T; 0 U; 15 Other;

Alignment Scores:

Pred. No.:	2,298-192	Length:	3998
Score:	2045.50	Matches:	396
Percent Similarity:	94.76%	Conservative:	2
Best Local Similarity:	94.29%	Mismatches:	21
Query Match:	92.72%	Indels:	1
DB:	12	Gaps:	1

US-10-798-532-2 (1-419) x ADQ24401 (1-3998)

QY 1 MetAlaGlySerGlyCysAlaATrPGlyAlaGluProProArpPheLeuGluAlaPheGly 20
 DB 160 ATGGCCGAGATCCCGCTGCGCTGCGCGCGAGCCCGCGTTTCTGAGAGCGCTTCGCG 219
 QY 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
 DB 220 CGCTGTGGCAGGTACAGAGCCCTGTGGTACGCGCTCTCCCGCTCGGTATCCGGTT 279
 QY 41 ArgCysArgGlyValAnProGlySerProProGlyValAlaLeuGlnPheLeuProProGly 60
 DB 280 CGCTGTGGCAGGTACAGAGCCCTGTGGTACGCGCTCTCCCGCTCGGTATCCGGTT 339
 QY 61 ThrThGlyValAlaAlaSerAlaAlaGluTyrGlyPheArgGlySerGluArgAlaAlaLeu 80
 DB 340 ACCACCGGGGCTGCGCTGCTGCGCGAGATGTTTAAAGCATGTTCTTACCAAGCG 399
 QY 81 GluGlnLeuGlnGlyHisArgAenIleValThrLeuTyrGlyValPheThrIleHis--- 99
 DB 400 AAGTCTGTGTTACCTTAAACCGAGAGTCACTTGTATGAGATGTTTAAACCGAGTT 459
 QY 100 PheSerProAnValProSerArgCysLeuLeuLeuGlnLeuLeuAlaSerValSer 119
 DB 460 GTTCTCAAGTGTGCATACCGCTGTGTTCTTGAATCTCGGATGTCAGTGTTCG 519
 QY 120 GluLeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAla 139
 DB 520 GAATTCCTCTTAATTCACATCCAGAGTGTTCATGTGATGATGATGATGATGATGATG 579
 QY 140 ArgAspValLeuGlnAlaLeuAlaPheLeuHisIleGluGlyTyrValHisAlaAspLeu 159
 DB 580 CGAGATGTTTGGAGGCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 639
 QY 160 LysProArgAnIleLeuTyrSerArgGluAnGluCysPheLeuLeuIleAspPheGly 179
 DB 640 AAACCAAGTAAATATGTCAGATGTCAGAGATGAATGATGATGATGATGATGATGATG 699
 QY 180 LeuSerPheLeuGlnGlyAnGlnAnPheValTyrIleGlnThrArgGlyTyrArgAla 199
 DB 700 CTTAGCTTCAAAAGAGCATCAGATGTAAGTAAATTAATTCAGACAGCGGATTCGCGCT 759
 QY 200 ProGlnAlaGluLeuGlnAnGlnCysLeuAlaGlnAlaGlyLeuGlnAnSerArgTrpGluCys 219
 DB 760 CCGAAGCAGAAATTCGAAATTCCTTGGCCAGGCTGCTGAGATGATACAGATG 819
 QY 220 ThrSerAlaValAspLeuTyrSerLeuGlyIleIleLeuLeuGlnLeuPheSerGlyMet 239
 DB 820 ACCTCAGCTGTGATCTGTGAGACCTTGAATCATTTTACTGAAATGTTCTCAGAAATG 879
 QY 240 LysLeuLysHisThrValArgSerGlnGluTrpLysAlaAnSerSerAlaIleIleAsp 259
 DB 880 AAATGAAATATCAGTCAGATCTCAGAAATGAAAGGAAAGGATTCGCTATTAATGAT 939
 QY 260 HisIlePheAlaSerLysAlaValAlaAlaIleProAlaTyrHisLeuArgAsp 279

DB 940 CACATATTTGCCAGTAAAGCATGTGTAATGCCGCAATTTCCAGCTTATCAGCTTAAGAGC 999
 QY 280 LeuIleLysSerMetLeuHisAspArpProSerArgGlyIleProAlaGluMetAlaLeu 299
 DB 1000 CTATTCAAAAGATCTTCATATATATCCAGACAGAAATTCCTGCTGAATGGCATTTG 1059
 QY 300 CysSerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuPro 319
 DB 1060 TGCAGCCCATTTTTCAGATTCCTTTGCCCCCTCATATTTGAAGATCTGGTCACTGCC 1119
 QY 320 ThrProValLeuArgLeuLeuAnValLeuAspAspArgTyrLeuGlyAnGlnGluGlu 339
 DB 1120 ACTCCAGTCTTAAGATCTGTAATGTGCTGATGTATGATTTCTTGAATGAAAGAGGA 1179
 QY 340 TyrGluAspValValGluAspValIleGlnGluGluCysGlnLysTyrGlyProValAlaSer 359
 DB 1180 TATGAAGATGTTGTGAAGATGTAAGAGGATGTCAAAAATATGACCAAGTGTATCT 1239
 QY 360 LeuLeuValProLysGlyLysAnProGlyArgGlyGlnValPheValGluTyrAlaAnAla 379
 DB 1240 CTACTTGTTCCAAAGAAATCTGCGACAGACAAAGTCTTGTGATGATCCAAATGCT 1299
 QY 380 GlysPheSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheVal 399
 DB 1300 GGTGATTCCAAAGCTGCGCAGAAATTAAGTGAAGGATGTTGATGGGAAATTTGTT 1359
 QY 400 ValAlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
 DB 1360 GTGGCTACATTTCTACCCGCTGATGTCCTAACAGAGGATATCTGATCAAACTTGTCT 1419
 RESULT 12
 AAH70764
 ID AAH70764 standard; cDNA; 529 BP.
 AC AAH70764;
 XX 19-SEP-2001 (first entry)
 DT 19-SEP-2001 (first entry)
 XX Human cervical cancer marker nucleic acid 2038.
 DE Human cervical cancer marker nucleic acid 2038.
 KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
 OS Homo sapiens.
 XX MO200142467-A2.
 PN MO200142467-A2.
 XX 14-JUN-2001.
 PD 14-JUN-2001.
 XX 08-DEC-2000; 2000WO-US033312.
 PF 08-DEC-1999; 99US-0169681P.
 XX 08-DEC-1999; 99US-0171350P.
 PR 21-DEC-1999; 2000US-0189315P.
 PR 14-MAR-2000; 2000US-0203791P.
 PR 12-MAY-2000; 2000US-0203791P.
 PR 09-JUN-2000; 2000US-0210600P.
 PR 21-JUL-2000; 2000US-0220114P.
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX Schlegel R, Deede J, Berger A, Zhao X;
 PI MPI, 2001-375006/39.
 DR MPI, 2001-375006/39.
 XX New isolated nucleic acid for diagnosing and treating cervical cancer and
 PT for assessing and detecting compounds for treating the cancer.
 XX Claim 1; Page 434; 1051p; English.
 PS The invention relates to novel genes (AAH68727-AAH73383) associated with
 CC cervical cancer with cytostatic activity. The nucleic acids and encoded
 CC polypeptides are useful: to assess if a patient is afflicted with
 CC cervical cancer or has a pre-malignant condition; to monitor the
 CC progression of cervical cancer or a premalignant condition in a patient;

CC and to select and/or assess the efficacy of a compound or therapy for
 CC inhibiting cervical cancer in a patient. The nucleic acids may also be
 CC useful for gene therapy

Sequence 529 BP; 156 A; 102 C; 132 G; 139 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	8,45e-69	Length:	529
Score:	789.00	Matches:	159
Percent Similarity:	95.78%	Conservative:	0
Best Local Similarity:	95.78%	Mismatches:	5
Query Match:	35.77%	Indels:	2
DB:	4	Gaps:	0

US-10-798-532-2 (1-419) x AAH70764 (1-529)

```

QY 220 ThSerAlaValAspLeuTPSPserLeuGlyIleIleLeuGluMetPheSerGlyMet 239
DB 26 ACCTCAGCTGTTGATCTGGAGACCTAGGAATCATTTTACTGGAATGTTCTCAGGAATG 85
QY 240 LyeLeuIleHisThrValArgSerGlnIleTrpIleValAsnSerSerAlaIleIleAsp 259
DB 86 AAACGAAACATACAGTCAGATCTCAGGAATGGAAGCAAAAGCTTCTGCTATTATTGAT 145
QY 260 HisIlePheAlaSerIleValAlaValAsnAlaIleProAlaTrpHisLeuArgAsp 279
DB 146 CACATATTTTCCAGTAAGCAGTGGTAATGCCGCAATTCACGCTATACCTTAAGAGAC 205
QY 280 LeuIleIleSerMetLeuHisAspAspProSerArgIleProAlaGluMetAlaLeu 299
DB 206 CTATATCAAAAGCATCTTCATGATGATCCAAAGAGAAGAAATTCCTGTAATGGCATTTG 265
QY 300 CysSerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuPro 319
DB 266 TGAAGCCCATCTTTAGCATCTCTTTGCCCTCATTTTGAATCTGTCATGCTTCC 325
QY 320 ThrProValIleuArgLeuAsnValLeuAspAspPheTrpLeuGluAsnGluGlu 339
DB 326 ACTCCAGTGTAGTACGCTGGAATGTGCTGATGATCATATTATCTTGAATGAAGAGAA 385
QY 340 TyrGluAspValValGluAspValIleGluGluGluGluGluGluGluGluGluGlu 359
DB 386 TATGAAGATTTGTTGTAAGATGTAAAGAGAGAGTGTCAAAATATGACCCAGGGGTAT 445
QY 359 GlnLeuValIleProGluAsnProGlyArgGlyGluValPheValGluTrpAlaAsn 379
DB 446 CTCTACTTGTCCAAAGGAAATCTCGGCAAGAAACAGCTTTGTTGATGCAAAAG 505
QY 379 IAGIYAspSerIleYs 383
DB 506 CTGGGATTCAAAG 519

```

RESULT 13

AAA02535
 ID AAA02535 standard; cDNA; 722 BP.

AC AAA02535;

DT 19-MAY-2000 (first entry)

XX Human colon cancer cell line polynucleotide sequence SEQ ID NO:2526.

XX Human; colon cancer; tumour; diagnosis; gene expression product; probe;

XX detection; cancerous state; metastasis; identification; breast cancer;

XX oestrogen receptor-positive breast cancer; therapy;

XX oestrogen receptor-negative breast cancer; lung cancer; ss.

OS Homo sapiens.

XX MO9958675-A2.

XX 18-NOV-1999.

PF 13-MAY-1999; 99WO-US010602.
 XX
 PR 14-MAY-1998; 98US-0085426P.
 PR 15-MAY-1998; 98US-0085537P.
 PR 15-MAY-1998; 98US-0085656P.
 PR 21-OCT-1998; 98US-0105234P.
 PR 27-OCT-1998; 98US-0105877P.

XX (CHIR) CHIRON CORP.
 PA (HYSE-) HYSEQ INC.

XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
 PI Reinhard C, Gleese K, Randazzo F, Kennedy GC, Fot D, Kassam A;
 PI Lamson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
 PI Leschowitz D, Kita D, Garcia V, Jones LW, Stache-Grain B;
 XX

DR WPI; 2000-126369/11.

PT Polynucleotide library used to determine cancerous states of mammalian
 cells.

PS Claim 1; Page 1020; 1097P; English.

XX AAA00010 to AAA02716 represent polynucleotides isolated from cDNA
 CC libraries constructed from human colon cancer cell lines. The present
 CC invention also describes a method of detecting differentially expressed
 CC genes correlated with a cancerous state of a mammalian cell, comprising
 CC detecting at least one differentially expressed gene product in a test
 CC sample derived from a cell suspected of being cancerous, where detection
 CC of the differentially expressed gene product is correlated with a
 CC cancerous state of the cell from which the test sample was derived. The
 CC polynucleotides sequences can be used in a method for detecting
 CC differentially expressed genes correlated with a cancerous state of a
 CC mammalian cell. The polynucleotides can also be used as probes for
 CC detecting and mapping related genes. They can be used in diagnosis and
 CC prognosis of diseases and disorders (e.g. identification of pre-
 CC metastatic or metastatic cancerous states, stages of cancer, or
 CC responsiveness of cancer to therapy). This is particularly for breast
 CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
 CC negative breast cancer, lung cancer, and colon cancer

XX Sequence 722 BP; 215 A; 129 C; 164 G; 193 T; 0 U; 21 Other;

Alignment Scores:

Pred. No.:	2,29e-66	Length:	722
Score:	766.50	Matches:	182
Percent Similarity:	84.86%	Conservative:	3
Best Local Similarity:	83.49%	Mismatches:	20
Query Match:	34.75%	Indels:	14
DB:	3	Gaps:	1

US-10-798-532-2 (1-419) x AAA02535 (1-722)

```

QY 152 GUGIYTYrValHiaAlaAspLeuIleProAlaGlnIleLeuTrpSerAlaGluAsnGlu 171
DB 1 GAGGGCTATGTCACGCGGCTCAAAACCACTAATATTTGAGAGTCAGAGATGA 60
QY 172 CysPheIleLeuIleAspPheGlyLeuSerPheIleGluGlyAsnGlnAspValIleTrp 191
DB 61 TGTTTTAACTCATTTGACTTTGACTTANCTTCAAGAGGCAATCAGATTTAAAGTAT 120
QY 192 IlegInThrAspGlyTYrArgAlaProGluAlaGluLeuGlnAsnCysLeuAlaGlnAla 211
DB 121 ATTCAAGACAGACGGGATCGGCTCCAGAAACA-GAATTGCAAAATTTGCTTGCCCANCT 179
QY 212 GylLeuGlnSerAspThrGlyCysThrSerAlaValAspLeuTrpSerLeuGlyIleIle 231
DB 180 GGCTTCAGAGATGATACAAATGTACTTCAGCTGTGATCTGTGAGAGCTTAGAAATCAT 239
QY 232 LeuLeuGluMetPheSerGlyMetCysLeuIleYsHleThrValArgSerGlnIleTrpIleYs 251
DB 240 TTACTGGAATATTTCTCAGGAATGAATCTGAATCATACAGTACAGATCTCAGGAATGGAAG 299

```


QY 252 AlaAsnSer-SerAlaIleIle-AspHisIlePheAlaSerIleValValAsnAl 270
 Db 300 GCAACAGTTTCTGCTATTATTGATCATTATTTCCAGTAAACCAANTGGTGAATGC 359
 QY 270 aaIaIleProAlaTyHisIleuArgAspLeuIlelySerMetLeuHisAspAsp-ProS 290
 Db 360 CCGAATTCACGCTATCACCTAANAGACCTTATCAAAAGATGCTTCATGATCATCCAA 419
 QY 290 eArg-ArgIleProAlaGluMetAlaIleu-CysSerProPhePheSer-IleProPheA 309
 Db 420 GCAAGAAAGATTCTCTGCTAAATGGCATTTGTCACCACTTCTTAAAGCATTCCTTTG 479
 QY 309 IaProHisIleGluAspLeuValMetLeuProThrProVal-IleuArgLeuAsnVal 328
 Db 480 CCCCTATATTGAAGATCTGATCATCTTCCACTCAGGCTAAGACTGCTGAATGTG 539
 QY 329 Leu-AspAspAspTyTleuGlyValAsnGlu-GluGluTyTleuAspValValGluAspVal 347
 Db 540 CTGGGATGATGATTATCTTGAGAAATGAAGAAGATTAAGAAGATGCTGTGAAGATGAT 599
 QY 348 TyGluGluCysGlnLys--TyTleuPro 356
 Db 600 AAAAGAAGAGTGGCAAAATNTTGCACCA 629

RESULT 14

AAH71966

ID AAH71966 standard; cDNA; 470 BP.

AC AAH71966;

DT 19-SEP-2001 (first entry)

DE Human cervical cancer marker nucleic acid 3240.

XX Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

XX Homo sapiens.

XX MO200142467-A2.

XX 14-JUN-2001.

XX 08-DEC-2000; 2000MO-US033312.

XX PF 08-DEC-1999; 99US-0169681P.

XX PR 21-DEC-1999; 99US-0171350P.

XX PR 14-MAR-2000; 2000US-0189315P.

XX PR 12-MAY-2000; 2000US-0203791P.

XX PR 09-JUN-2000; 2000US-0210600P.

XX PR 21-JUL-2000; 2000US-0220114P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Deeds J, Berger A, Zhao X;

XX DR WPI; 2001-375006/39.

XX PT New isolated nucleic acid for diagnosing and treating cervical cancer and

XX PS for assessing and detecting compounds for treating the cancer.

XX Claim 1; Page 636; 1051pp; English.

XX CC The invention relates to novel genes (AAH68727-AAH73383) associated with

XX CC cervical cancer with cytostatic activity. The nucleic acids and encoded

XX CC polypeptides are useful: to assess if a patient is afflicted with

XX CC cervical cancer or has a pre-malignant condition; to monitor the

XX CC progression of cervical cancer or a premalignant condition in a patient;

XX CC and to select and/or assess the efficacy of a compound or therapy for

XX CC inhibiting cervical cancer in a patient. The nucleic acids may also be

XX CC useful for gene therapy

XX CC Sequence 470 BP; 140 A; 91 C; 114 G; 125 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	6,51e-63	Length:	470
Score:	729.00	Matches:	147
Percent Similarity:	96.71%	Conservative:	0
Best Local Similarity:	96.71%	Mismatches:	3
Query Match:	33.05%	Indels:	2
DB:	4	Gaps:	0

US-10-798-532-2 (1-419) x AAH71966 (1-470)

QY 220 ThrSerAlaValAspLeuTrpSerLeuGlyIleIleLeuGluMetPheSerGlyMet 239
 Db 18 ACCTCAGCTGTTGATCTGTGGAGCTTAAAGATCATTTTAACTGGAATGTTCTCAGAAATG 77
 QY 240 TySLeuTyHisThrValArgSerGlnGluTlPlyAlaAsnSerSerAlaIleIleAsp 259
 Db 78 AAACGAAACATACGATCAGATCTCAGAAATGGAAGCAACAGTTCGCTATTATTGAT 137
 QY 260 HisIlePheAlaSerIleAlaValValAsnAlaIleProAlaTyHisIleuArgAsp 279
 Db 138 CACATATTTCGCGATGAAGCACTGTAATGCCGAAATTCACGCTTATCAGCTTAAGAGAC 197
 QY 280 LeuIleTySerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaIleu 299
 Db 198 CTTATCAAAAGCATCTTCATGATGATCCAAAGCAAGAAATTCGCTGAATGCAATTG 257
 QY 300 CysSerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuPro 319
 Db 258 TGCAGCCCATTCCTTAAAGATTCCTTTGCCCCCTCATATTGAAGATCTGTCATGCTTCCC 317
 QY 320 ThrProValLeuArgLeuLeuAsnValLeuAspAspAspTyTleuGlyValAsnGluGlu 339
 Db 318 ACTCAAGTCTTAAGACTGCTGTAATGTGCTGATGATGATTATCTTGAGAAATGAAGAGA 377
 QY 340 TyTleuAsp-ValValGluAspValIleGlu-GluCysGlnLysTyTleuProValValS 359
 Db 378 TATGAAGATGTTGTATGAAGATGTAAAGAGGAGTGTCAAAATATATGACCAAGGGATAT 437
 QY 359 eTleuLeuValProIleGluAsnProGlyArg 369
 Db 438 CTCTACTGTGTCAAAGAAATCCTGCGACA 469

RESULT 15

AAH69101

ID AAH69101 standard; cDNA; 461 BP.

AC AAH69101;

DT 19-SEP-2001 (first entry)

DE Human cervical cancer marker nucleic acid 375.

XX Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

XX Homo sapiens.

XX MO200142467-A2.

XX 14-JUN-2001.

XX 08-DEC-2000; 2000MO-US033312.

XX PF 08-DEC-1999; 99US-0169681P.

XX PR 21-DEC-1999; 99US-0171350P.

XX PR 14-MAR-2000; 2000US-0189315P.

XX PR 12-MAY-2000; 2000US-0203791P.

XX PR 09-JUN-2000; 2000US-0210600P.

XX PR 21-JUL-2000; 2000US-0220114P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Deeds J, Berger A, Zhao X;

DR WPI; 2001-375006/39.

XX New isolated nucleic acid for diagnosing and treating cervical cancer and
PT for assessing and detecting compounds for treating the cancer.

XX Claim 1; Page 168-169; 1051pp; English.

XX The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy

XX Sequence 461 BP; 134 A; 91 C; 109 G; 127 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.le-53	Length:	461
Score:	636.00	Matches:	138
Percent Similarity:	93.33%	Conservative:	2
Best Local Similarity:	92.00%	Mismatches:	5
Query Match:	28.83%	Indels:	6
DB:	4	Gaps:	0

US-10-798-532-2 (1-419) x AAH69101 (1-461)

```
QY 220 ThSerAlaValAspLeuTppSerLeuGlylleleuLeuGluMetPheSerGlyMet 239
    |||||
DB 22 ACCTCAGCTGTTATCTGTGAGCTTGAAGCATTTTACTGAAAGTTCTCAGAAATG 81
    |||||
QY 240 LysLeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleLeuAsp 259
    |||||
DB 82 AAACGAAACATACAGTCAGATCTCAGGAATGGAAGCAACAGTCTGCTATTATTGAT 141
    |||||
QY 260 HisIlePheAlaSerLysAlaValAlaAsnAlaIleProAlaTyrHisLeuArgAsp 279
    |||||
DB 142 CACATATTGCCAGTAAGCAGTGTGAATGCCGCAATTCAGCCTATCCCTAAGAGAC 201
    |||||
QY 280 LeuIleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeu 299
    |||||
DB 202 CTATACAAAGCATGCTTCATGATGATCCAAAGCAAGAAATTCCTGCTGAATGGCATTG 261
    |||||
QY 300 CysSerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuPro 319
    |||||
DB 262 TGCAGCCCATCTTTTGACATCTCTTTGCCCTCATAT-GAAGATCTGGTCATGCTTCCC 320
    |||||
QY 320 ThrProValLeuArgLeuLeuAsnValLeuAspAspTyrLeuGlyAsnGluGlu 339
    |||||
DB 321 ACTCCAGTGTGAAGATGCTGAATGCTGATGATGATATCTTGAGATGAAGAGAGAA 380
    |||||
QY 340 TyrGluAsp--ValValGluAspValLysGlu--GluCysGlnLysTyrGlyPro--ValY 358
    |||||
DB 381 TATGAAGATTTGTTAAGAAATGAAGAGGAGGTGTCAAAAATATGACCAAGGTGG 440
    |||||
QY 358 aLserLeuValProLys 364
    |||||
DB 441 TATCTCTACTTGTGCCAA 460
    |||||
```

Search completed: November 29, 2004, 11:42:18
Job time : 521 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 29, 2004, 11:25:02 ; Search time 3159.5 Seconds
(without alignments)
4832.489 Million cell updates/sec

Title: US-10-798-532-2

Perfect score: 2206
Sequence: 1 MAGSGCAGCABPPRFLEARG.....VATFVPLSAKRGYLYQTLL 419

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+.p2n.model -DEV=xlp
-Q=/cgn2.1/USPTO.epool.p/US10798532/runat_29112004_112453_15901/app_query.faeta_1.1166
-DB=EST -QFMT=fastap -SUFFIX=rat -MIMATCH=0.1 -LOOPEXT=0 -LIST=45
-UNITS=bites -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi -LST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORC=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10798532_@CCN_1_1_6628_@runat_29112004_112453_15901 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:.*
1: gb_est1:.*
2: gb_est2:.*
3: gb_hic:.*
4: gb_est3:.*
5: gb_est4:.*
6: gb_est5:.*
7: gb_est6:.*
8: gb_gse1:.*
9: gb_gse2:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2174	98.5	1763	3	AK013347 Mus muscu
2	1711	77.6	2846	3	AK030152 Mus muscu
3	1521.5	69.0	1079	4	BM468107 AGENCOURT
4	1274.5	57.8	1015	6	BY712344
5	1242	56.3	917	4	BM451184 AGENCOURT
6	1170	53.0	823	7	CO771206 testis ES
7	1162	52.7	779	5	BU447831 603767571
8	1159	52.5	896	7	CO773091 testis ES
9	1133	51.4	781	7	CN232453 WtB074D11

10	1110	50.3	932	5	BU541306	BU541306 AGENCOURT
11	1106	50.1	717	5	BU85761	BU85761 603582650
12	1091	49.5	899	2	BE895119	BE895119 601436041
13	1079	48.9	772	4	BT62420	BT62420 60326316
14	1075	48.7	794	5	BU205357	BU205357 603104559
15	1073.5	48.7	834	7	CF593363	CF593363 AGENCOURT
16	1068.5	48.4	885	4	BI259568	BI259568 602968330
17	1066	48.3	875	4	BQ225071	BQ225071 AGENCOURT
18	1053	47.7	682	4	BI333581	BI333581 602997069
19	1050	47.6	657	4	BG547675	BG547675 602575520
20	1028	46.6	844	4	BG474616	BG474616 602517423
21	1014	46.0	802	7	CO737403	CO737403 SLT02c17
22	1008.5	45.7	876	5	BU184346	BU184346 AGENCOURT
23	990.5	44.9	906	4	BG386905	BG386905 602454704
24	968.5	43.9	966	5	BU39278	BU39278 604167528
25	967	43.8	631	5	BU112717	BU112717 603131421
26	966.5	43.8	1003	5	BQ922020	BQ922020 AGENCOURT
27	962	43.6	755	5	BU371858	BU371858 603568233
28	958	43.4	699	5	BU070784	BU070784 UI-M-FRO-
29	944	42.8	604	5	BX261107	BX261107
30	909.5	41.2	716	7	CF745774	CF745774 UI-M-GVO-
31	900	40.8	570	5	BU293329	BU293329 604166812
32	896	40.6	670	7	CN359673	CN359673 170004245
33	892	40.4	928	5	BQ424438	BQ424438 AGENCOURT
34	880.5	39.9	824	2	BE568528	BE568528 601342425
35	870.5	39.5	820	7	CK597557	CK597557 AGENCOURT
36	862	39.1	672	5	BX261106	BX261106
37	861.5	39.1	1063	6	CD510246	CD510246 CDA100-G0
38	858	38.9	785	6	CB990192	CB990192 AGENCOURT
39	850	38.5	523	3	BF929410	BF929410 IL2-NT020
40	828.5	37.6	1063	5	BU356643	BU356643 603474124
41	818.5	37.1	743	4	BI562440	BI562440 603256324
42	818	37.1	763	7	CK364431	CK364431 AGENCOURT
43	801	36.3	933	2	BE886040	BE886040 601507192
44	796	36.1	986	5	BU751496	BU751496 CH4#002_E
45	793	35.9	484	5	BX281211	BX281211 BX281211

ALIGNMENTS

RESULT 1
AK013347 LOCUS
DEFINITION Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length enriched library, clone:281045604 product:kinase interacting with leukemia-associated gene (etractmin), full insert sequence.

ACCESSION AK013347
VERSION AK013347.1 GI:12850651
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Fukayocita, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 Carninci, P., Shibata, Y., Hayata, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)
2 MEDLINE 99279253
PUBMED 10349636

REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayata, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)

REFERENCE
AUTHORS Shihata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kutsuna, T., Tahito, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujitake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,

Db	939	ATATTGCCAGTAAACAGTGGATGCGCAATTCACGCTATCATCCAGACCTT	998
Qy	281	IIeIySeMeIeuthiSaspAppProSerArXgIleProIlaIuMeIaIaIuCyS	300
Db	999	ATCAAAAGCATGTTCAATGATGACCCGGGCAAGAATCCCTGCTAGATGGCATGTGC	1058
Qy	301	SePProPheSerIleProPheAaIProHilGluIuSpIeuValMeIeUpProthr	320
Db	1059	AGCCCATCTTTTGCAATTCCTTTTGCCCTCATATTTGAAGATTCGTGATGCTCCGACT	1118
Qy	321	ProValIeuArgIeuIeuIeuValIeuAspAspAspTyrIeuGluIuSengIuIuYr	340
Db	1119	CCAGTCTCAGGCTCCCTCATATGCTGATGATGATTAATCTGAMAATGAAATGATATAT	1178
Qy	341	GIuAspValValGIuIuSpValIyGIuIuGIuIuYrIyIyIyProValIaISeIeu	360
Db	1179	GAAGATGTTGTAAAGATGGAAGAGAGATGTCAGAAATATGACCCAGTCTTCTTG	1238
Qy	361	IeuValProIyGIuIuSpProGIyArGIyGIuIuValIyIyIyIyIyIyIyIyIyIyIy	380
Db	1239	CTTGTTCCAAAGAAATCTTGCGACAGGCAAGTCTTGATGATGCGGAACGCTGGT	1298
Qy	381	AspSerIyIyAaIaIaGIuIyIeIuIyIeIuIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy	400
Db	1299	GATTCCAAAGCTGCTCAGAAATGTTGCTGACCTGGAGGATGTTGACGGGAAGTTGTTGTC	1358
Qy	401	AlIeThPheIy	419
Db	1359	GCTACATTTTACCCGCTGAGTGCCTACAGAGGAGATCTTATCAACCTTGCTT	1415
RESULT 2			
LOCUS	AK030152	2846 bp	mRNA linear HTC 03-APR-2004
DEFINITION			Mus musculus adult male testis cDNA, RIKEN full-length enriched
ACCESSION			library, clone:4933401D07 product:kinase interacting with
KEYWORDS			leukemia-associated gene (strathmin), full insert sequence.
VERSION			AK030152
ORGANISM			AK030152.1 GI:26326136
SOURCE			HTC; CAP trapper.
ORGANISM			Mus musculus (house mouse)
REFERENCE			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS			Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE			1 Carninci, P. and Hayashizaki, Y.
AUTHORS			High-efficiency full-length cDNA cloning
JOURNAL			Meth. Enzymol. 303, 19-44 (1999)
MEDLINE			99279253
PUBMED			10348636
REFERENCE			2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
AUTHORS			Itoh, M., Komano, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE			Normalization and subtraction of cap-trapper-selected cDNAs to
JOURNAL			prepare full-length cDNA libraries for rapid discovery of new genes
MEDLINE			Genome Res. 10 (10), 1617-1630 (2000)
PUBMED			20499374
REFERENCE			3
AUTHORS			Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
TITLE			Komano, H., Akiyama, J., Nishii, K., Katsunari, T., Tashtiro, H., Itoh, M.,
JOURNAL			Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
MEDLINE			Yamamoto, R., Katsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
PUBMED			Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohata, E., Watabiki, M.,
REFERENCE			Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J.,
AUTHORS			Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE			RIKEN integrated sequence analysis (RISA) system--384-format
JOURNAL			sequencing pipeline with 384 multicapillary sequencer
MEDLINE			Genome Res. 10 (11), 1757-1771 (2000)
PUBMED			20530913
REFERENCE			4
AUTHORS			The RIKEN Genome Exploration Research Group Phase II Team and the

TITLE	PANTOM Consortium.
JOURNAL	Functional annotation of a full-length mouse cDNA collection
REFERENCE	Nature 409, 688-690 (2001)
AUTHORS	5
TITLE	The PANTOM Consortium and the RIKEN Genome Exploration Research
JOURNAL	Group Phase I & II Team.
MEDLINE	Analysis of the mouse transcriptome based on functional annotation
PUBMED	of 60,770 full-length cDNAs
REFERENCE	Nature 420, 563-573 (2002)
AUTHORS	6 (bases 1 to 2846)
TITLE	Adechi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
JOURNAL	Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M.,
MEDLINE	Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T.,
PUBMED	Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
REFERENCE	Kato, H., Kawai, J., Kojima, Y., Kondo, S., Komano, H., Kouda, M.,
AUTHORS	Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
TITLE	Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Ohata, N.,
JOURNAL	Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
MEDLINE	Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
PUBMED	Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
REFERENCE	Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
AUTHORS	Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
MEDLINE	Physical and Chemical Research (RIKEN), Laboratory for Genome
PUBMED	Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
REFERENCE	RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
AUTHORS	Kanagawa 230-0045, Japan (E-mail:genome-res@gscc.riken.jp,
TITLE	URL:htp://genome.gsc.riken.jp/, Tel:81-45-503-9222,
JOURNAL	Fax:81-45-503-9216)
MEDLINE	cDNA library was prepared and sequenced in Mouse Genome
PUBMED	Encyclopedia Project of Genome Exploration Research Group in Riken
REFERENCE	Genomic Sciences Center and Genome Science Laboratory in RIKEN.
AUTHORS	Division of Experimental Animal Research in Riken contributed to
TITLE	prepare mouse tissues.
JOURNAL	Please visit our web site for further details.
MEDLINE	URL:htp://genome.gsc.riken.jp/
PUBMED	URL:htp://fantom.gsc.riken.jp/.
REFERENCE	Location/Qualifiers
AUTHORS	1..2846
TITLE	/organism="Mus musculus"
JOURNAL	/mol_type="mRNA"
MEDLINE	/strain="C57BL/6J"
PUBMED	/db_xref="FANTOM DB:4933401D07"
REFERENCE	/db_xref="taxon:10090"
AUTHORS	/clone="4933401D07"
TITLE	/sex="male"
JOURNAL	/cissue_type="testis"
MEDLINE	/clone_lib="RIKEN full-length enriched mouse cDNA library"
PUBMED	/dev_stage="adult"
REFERENCE	238..1230
AUTHORS	/note="unnamed protein product; kinase interacting with
TITLE	leukemia-associated gene (strathmin) (MGI:1341908)
JOURNAL	putative"
MEDLINE	/codon_start=1
PUBMED	/protein_id="BAC26812.1"
REFERENCE	/db_xref="GI:26326137"
AUTHORS	/translation="MTLGVFTIHSFNVPSKCLLEIDVSELLYSXSGSGSM
TITLE	MIQCARVLEALAFHHEGYHADLPKNITLWSAENEFKIDFGLSFKENQDVKY
JOURNAL	IOTDGRAPAEALQNCIAQGLQSDTECSAVDLISGIILEMSSGMLKRTVSQF
MEDLINE	WKANSASIDHIFASKAVVNAIPIYALIDILKSMIHDGPRIPAEMLCSPPFSIP
PUBMED	PAPHIEDIWMLEPTVILRLNVDDIYLEDDEEDYEVVEDECKQYGVSVLWKE
REFERENCE	NGRGQVPEVYVNAEDSKAQCLTLGRMGDFVATFPLASAYRGYIYQTL"
AUTHORS	2823..2828
TITLE	/note="putative"
JOURNAL	2846
MEDLINE	/note="putative"
PUBMED	2846
REFERENCE	/note="putative"
AUTHORS	Alignment Scores:
TITLE	Pred. No.:
JOURNAL	Score:
MEDLINE	Percent Similarity:
PUBMED	2.78e-168
REFERENCE	Length:
AUTHORS	1711.00
TITLE	Matches:
JOURNAL	326
MEDLINE	Conservative:
PUBMED	2
REFERENCE	
AUTHORS	

```

Beet Local Similarity: 98.79%      Mismatches: 2
Query Match: 77.56%      Indels: 0
DB: 3      Gaps: 0
US-10-798-532-2 (1-419) x AK030152 (1-2846)

QY 90 valThrlentyrGlyValPheThrIleHisPheSerProAsnValProSerArgCysLeu 109
   :::
Db 238 ATGACTTATATGAGAGCTTATACATACACTTCTCTCCCATGAGCATACAGCTGCTG
QY 110 LeuLeuGluLeuLeuAspValSerValSerGluLeuLeuLeuTyrSerSerHisGlnGly 129
Db 298 GCGCTTGAACTCTCGATGTCAGAGTTTCGAAATGCTTTTATATTCATCCATCATCAGGCG
QY 130 CysSerMetTrpMetIleGlnHisCysValArgAspValLeuGluAlaLeuAlaPheLeu 149
Db 358 TGGTTCATGATGATGATACAGCATGTCGCCGAGATGTTTGGAGAGCCCTCGCTTTCTT 417
QY 150 HisHisGluGlyTyrValHisAlaAspLeuLysProArgAsnIleLeuTPSerAlaGlu 169
Db 418 CACCATGAGGGCTACGTCATCGACACTCAACACAGAAACATCCTGTCGAGTCGAG 477
QY 170 AsnGluCysPheLysLeuIleAspPheGlyLeuSerPheLysGluGlyAsnGlnAspVal 189
Db 478 AATGAGGTTTAACTTATGACTTGACTTGAAGCTTCAAGAGGCAATCAGAGCGTA 537
QY 190 LysTyrIleGlnThrAspGlyTyrArgAlaProGluAlaGluLeuGlnAsnCysLeuAla 209
Db 538 AAGTATATTCAGACAGCGGGTATCGGCTCTCTGAGCAGAGCTGCGAACTGCTTGCC 597
QY 210 GlnAlaGlyLeuGlnSerAspThrGluCysThrSerAlaValAspLeuTPSerLeuGly 229
Db 598 CAGCGCGGCTGCGAGAGTATACAGAAATGACCTGAGCTTGATGTCGAGACCTCGGA 657
QY 230 IleIleLeuLeuGlnMetPheSerGlyMetLysLeuLysHisThrValArgSerGlnGlu 249
Db 658 ATCATTTTACTGAAAGTTCTCAGAGATAGAGCTGAAACATCAGACATCTCAGGAA 717
QY 250 TrpLysAlaAsnSerSerAlaIleIleAspHisIlePheAlaSerLysAlaValAsn 269
Db 718 TGGAGGCGAAACAGTCTGCTATTTATTTGATCATATATTTCCAGTAAAGCAGTGGTAAT 777
QY 270 AlaAlaIleProAlaTyrHisLeuArgAspLeuIleLysSerMetLeuHisAspAspPro 289
Db 778 GCGCGAATTCACACCTATCACTCAGAGACCTTATCAAAAGCATGCTTCATGATGACCG 837
QY 290 SerArgArgIleProAlaGluMetAlaLeuCysSerProPhePheSerIleProPheAla 309
Db 838 GGCAGAAAGATCCCTGCTGATGAGTGGCATGTGAGCCCATCTTTAGCATTCCTTTTGGC 897
QY 310 ProHisIleGluAspLeuValMetLeuProThrProValLeuArgLeuLeuAsnValLeu 329
Db 898 CCTCATATTTGAAGATCTGTCGATGCTTCGACATCCAGTGTCTCAGGCTCCCTCAATGTCG 957
QY 330 AspAspAspTyrLeuGlyAsnGluGluGlyTyrGluAspValValGluAspValLysGlu 349
Db 958 GATGATGATTAACCTGAAATGAAGATGATGAAGATGTTGAGAAAGATGGAAGAG 1017
QY 350 GluCysGlnLysTyrGlyProValIleSerLeuLeuValProLysGluAsnProGlyArg 369
Db 1018 GAGTGTCTAGAAATATGACAGAGTGTCTCTGCTGTGTTCCAAAGGAAATCTCTGCGAGA 1077
QY 370 GlyGlnValPheValGluTyrAlaAsnAlaGlyAspSerLysAlaAlaGlnLysLeu 389
Db 1078 GGAACAAGTCTTCTGTAGTACCGGAAGCTGTCGATTCCAAAGCTGTCGAAGAGTCTG 1137
QY 390 ThrGlyArgPheCysPheAspGlyLysPheValValAlaThrPheTyrProLeuSerAlaTyr 409
Db 1138 ACTGGAGAGATGTTTGAAGGAGTTTGTGTGGCTACATTCAACCGGTGAGTGCCTAC 1197
QY 410 LysArgGlyTyrLeuTyrGlnThrLeuLeu 419
Db 1198 AAGAGGGAGATATCTTATCAAACTTGTCTT 1227

```

```

RESULT 3
BM468107
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
CONTACT: Robert Strausberg, Ph.D.
Email: cga@bs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12224 row: b column: 06
High quality sequence stop: 754.
Location/Qualifiers
1..1079
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5535653"
/risue_type="leiomysarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/notes="Organ: uterus; Vector: pCMV-SPORT6, Site_1: NotI,
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb."
ORIGIN
Alignment Scores:
Pred. No.: 5,61e-149 Length: 1079
Score: 1521.50 Matches: 309
Percent Similarity: 92.86% Conservative: 3
Beet Local Similarity: 91.96% Mismatches: 17
Query Match: 68.97% Indels: 7
DB: 4 Gaps: 3
US-10-798-532-2 (1-419) x BM468107 (1-1079)
QY 90 valThrlentyrGlyValPheThrIleHisPheSerProAsnValProSerArgCysLeu 109
   |||
Db 2 GATCACTTGATGAGATGTTTACATCCACTTTCTCCAAATGTCATCAGCTGTCG 61
QY 110 LeuLeuGluLeuLeuAspValSerValSerGluLeuLeuTyrSerSerHisGlnGly 129
Db 62 TTGCTTGAACCTCTGATGATCAGTGTTCGAAATGCTCTTATATTCAGTCAACAGGCT 121
QY 130 CysSerMetTrpMetIleGlnHisCysValArgAspValLeuGluAlaLeuAlaPheLeu 149
Db 122 TGTTCATGATGATGATACAGCATGTCGCCAGAGATGTTTGGAGGCCCTTCTTTCTT 181
QY 150 HisHisGluGlyTyrValHisAlaAspLeuLysProArgAsnIleLeuTPSerAlaGlu 169
Db 182 CATCATGAGGGCTATGTCATGCGACCTCAACACAGTAACATATTGTCGAGTCGAGAG 241
QY 170 AsnGluCysPheLysLeuIleAspPheGlyLeuSerPheLysGluGlyAsnGlnAspVal 189
Db 242 AATGATGTTTAACTATGACTTGAAGCTTGAAGCTTCAAGAGGCAATCAGAGTGA 301
QY 190 LysTyrIleGlnThrAspGlyTyrArgAlaProGluAlaGluLeuGlnAsnCysLeuAla 209

```

DB 302 AGTATATTCAGACACACCGGTATCGGGCTCCGAGAGCAATTCGAAATTCCTGGCC 361

QY 210 GlnA1AGlyLeuGlnSerAspThrGluCysThrSerAlaValAspLeuTrpSerLeuGly 229

DB 362 CAGGCTGGGCTCGAGAGTATATGAGATGATACCTCAGCTGTTGATCTGTGAGCCCTAGGA 421

QY 230 IleIleLeuGlnGlnMetPheSerGlyMetLeuValLeuHisPheAlaArgSerGlnGlu 249

DB 422 ATCATTTTATCTGAGAAATGTTCTGAGAGAACTGAAACATACAGTCAGATTCAGAGAA 481

QY 250 TrpValAsnSerSerAlaIleIleAspHisIlePheAlaSerIleValAlaValAsn 269

DB 482 TGGAGAGCAACAGCTCTGCTATATTGATGATCAATTTCCGCTAAGAGAGTGAGAT 541

QY 270 AlaAlaIleProAlaTrpHisLeuAlaGlyAspLeuIleLeuSerMetLeuHisAspAspPro 289

DB 542 GCGGCAATTCAGACCTATCCCTTACAGAGACCTTATCAAAAGCATGCTTCATGATGATCA 601

QY 290 SerAspArgIleProAlaGluMetAlaLeuGlySerProPhePheSerIleProPheAla 309

DB 602 AGAGAGAGAAATTCCTCTGAGAAATGCAATGTCAGCCCATCTTTAGCATCTCTTTGCC 661

QY 310 ProHisIleGlnAspLeuValMetLeuProThrProValIleuArgLeuLeuValIleu 329

DB 662 CCTCATATTGAAGATCTGGTCACTCCACCTCCAGCTGCTAGACCTGCGAATGCTGCTG 721

QY 330 AspAspAspTrpLeuGlnAspGlnGluGluTrpGlnAspValAlaGlnAspValValGlu 349

DB 722 GATGATGATGATATCTTATGAGATGAGAGAAATGATGAGATGATGAGATGATGAGAG 781

QY 350 GlnGlyGlnLeuTrpGlyProValIleSerLeuValProGlyGluAspProGlyArg 369

DB 782 GAGTGTCAAAATATGACACCGAGTATCTCTGATCTCTGATCTCTGATCTCTGACAG 841

QY 370 GlnGlnValI-PheValGlnTrpAlaAsnAla-GlyAspSerIleValAlaGlnIleLeu 388

DB 842 GGAACAAGCTCTTTGATGAGATGCAATGCTGATGATGATGATGATGATGATGATGATGAT 901

QY 389 LeuThrGlnArg--MetPheAspGlyLeuPheValAlaAla-ThrPheTrpProLeu-- 406

DB 902 CTGACCTGAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 961

QY 407 -SerAlaTrpValAspArgGly---TyrLeuTrpGlnThrLeu 418

DB 962 GAGTGTCTTACAGAGAGGAGATTAATCTGATTCAAACCTT 1001

RESULT 4
LOCUS BY12344 1015 bp mRNA linear EST 17-DEC-2002
DEFINITION BY12344 RIKEN full-length enriched, 10, 11 days embryo whole body
VERSION BY12344
KEYWORDS BY12344.1 GI:27123621
SOURCE EST.
ORGANISM Mus musculus (house mouse)
MUS musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1015)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Otsu, N., Saito, R., Suzuki, H., Yamakawa, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Kiyosawa, H., Gojohori, T., Baldarelli, R., Hill, D.P., Bull, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Brad, D., Bruscia, V., Fletcher, C., Corbani, L.E., Cousins, S., Dalla, R., Dargatz, T.A., Gariboldi, M., Gisel, C., Godzik, A., Gough, J., Grimmond, S., Gustlinch, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltsev, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,

Numata, K., Okido, T., Pavan, W.J., Perce, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramchandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takemura, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wymshaw-Borle, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavalan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, U., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imocani, K., Ishii, Y., Itoh, M., Kogawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Landre, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
22354683
JOURNAL MEDLINE
PUBMED
COMMENT

Contract: Yoshinori Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-research.riken.jp, URL: http://genome.gsc.riken.jp/
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Horii, F., Imocani, K., Ishii, Y., Itoh, M., Kogawa, I., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Koye, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken.
Division of Experimental Animal Research in Riken contributed to prepare mouse cDNAs.
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
Location/Qualifiers
1..1015
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="2810454004"
/tissue.type="whole body"
/dev_stage="10, 11 days embryo"
/clone.lib="RIKEN full-length enriched, 10, 11 days embryo whole body"

FEATURES
source
ORIGIN
Alignment Scores:
Pred. No.: 4, 99e-123
Score: 1274.50
Percent Similarity: 89.51%
Best Local Similarity: 88.11%
Query Match: 57.77%
Length: 1015
Matches: 252
Conservative: 4
Mismatch: 23
Indels: 7
Gaps: 4

DB: 6
Gaps: 4

US-10-798-532-2 (1-419) x BY712344 (1-1015)

QY 1 MetAlAGlySerGlyCysAlaATPGlyAlaGluProProArphpheluGluAlaAphGly 20
 DB 159 ATGGCGGGGCTCCGGCTGGCGGGGCGGACCGCGGCTTCTCGAGGGCTTCGGG 218
 QY 21 ArgLeuTrpGlnAlaGlnSerArgLeuGlySerGlySerSerAlaSerValTYrArgVal 40
 DB 219 CGGCTGTGGAGGTCCAGACCGCCCTGGGAGCGGCTCGTCCGCTCGGTACCGGGT 278
 QY 41 ArgCysArgGlyAlaAnProGlySerProProGlyAlaLeuGlnAphLeuProProGly 60
 DB 279 CGGTGCTGGCGGTACCGCGGCTCGCCCGCGGCGCTCAAGCAGTTCCTCCGCGGGA 338
 QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTYrGlyPheArgGlyGluArgAlaAlaLeu 80
 DB 339 ACCACCGGGGCTCGGCTCGCGGAGTATGTTCCGCAAGAGAGGGCGGCGCTG 398
 QY 81 GluGlnLeuGlnGlyHisArgAlaAlaValThrLeuTYrGlyValPheThrIleHisPhe 100
 DB 399 GAGCAGTGGAGGGTCCAGAGACATGTCATTTATATGAGTCTTTACCATACACTTC 458
 QY 101 SerProAlaValProSerArgCysLeuLeuGluLeuLeuAlaSerValSerGly 120
 DB 459 TCTCCCAATGTGCATCAGCCTGTCTGCTGCTGAACTCCTGGATGTCAGTGTTCGGA 518
 QY 121 LeuLeuLeuTYrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
 DB 519 TTGCTTTATATTCAGTCAACAGGGCTGTCCATGTGATGATACAGCATGTGGCCGA 578
 QY 141 AspValLeuGlnAlaLeuAlaPheLeuHisGlnGlyTYrValHisAlaAspLeuVal 160
 DB 579 GAGTTCGTGGAGGCCCTCGCTTTCTTCAACATGAGGGCTACCGCATGACAGCTCAAA 638
 QY 161 ProArgAlaValLeuTrpSerAlaGluArgGlyCysPheLeuValLeuAspPheGlyLeu 180
 DB 639 CCACGAAACATCTCGAGAGTGGCGGAGATGAGTGTAAAGTTCATTTGACTTGGACTC 698
 QY 181 SerPheLeuGlnGlyAlaGlnAlaPheValLeuTYrIleGlnThrAspGlyTYrArgAlaPro 200
 DB 699 AGCTTCAAAAGCAATTCAGACGTAATATATTCAGACAGCGGCTATCCGCTCCT 758
 QY 201 GluAlaGluLeuGlnAlaGlnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGlyCysThr 220
 DB 759 GAGCAGAGCTGCAGAACTGCTTGGCCCAAGCCGCGCTGCAGATGATACAGATGTACC 818
 QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLeu 240
 DB 819 TCAGCTGTGATCTGTCCGGGCTCGGAGATCACTTACTGAGATGTTCTCAGGAATGAG 878
 QY 241 LeuLeu---HisThrValArgSerGlnGluTrpIleAlaAsnSerSerAlaIleIleAsp 259
 DB 879 CTAAATAATACATCAGATCTCAGCA-----TGGAGGCA---AGCACTTCGCATCACTTCG 929
 QY 260 HisIlePheAlaSerLeuValAlaAlaAlaIleProAlaTYrHisIleLeuArgAsp 279
 DB 930 CATATATTT-----GCAGTGAGCAGTGTGTGCAATTCAAGCATCACTCAGAC 980
 QY 280 LeuIleLeuSerMetLeu 285
 DB 981 GCTTATCAGAGCATGCTC 998

RESULT 5 BM451184 917 bp mRNA linear EST 05-FEB-2002
 LOCUS BM451184
 DEFINITION AGENTCOURT_6392719 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5495159
 5', mRNA sequence.
 ACCESSION BM451184
 VERSION BM451184.1 GI:18500224
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 1 (bases 1 to 917)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgaabs-r@mai.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM12121 row: j column: 24
 Location/Qualifiers
 High quality sequence stop: 718.

FEATURES
 source
 1..917
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5495159"
 /cissue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_1b="NIH MGC 67"
 /note="Organ: eye; Vector: pCMV-Sport6; site_1: NotI;
 site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.75 kb. Library constructed by Life
 Technologies."

ORIGIN

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
1,136-119	917	1242,00	240	0	1	0
Score:	99.598					
Percent Similarity:	99.598					
Best Local Similarity:	56.308					
Query Match:	4					

US-10-798-532-2 (1-419) x BM451184 (1-917)

QY 179 GlyLeuSerPheGlyGlnGlyAlaPheValTYrIleGlnThrAspGlyTYrArg 198
 DB 22 GGGCTTACCTTCAAGAGCAATCAGAGTGTAAATATTCAGACAGCGGATATCG 81
 QY 199 AlaproGluAlaGluLeuGlnAlaGlnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGlu 218
 DB 82 GCTCCAGAAAGCAATTCAGAAATGCTTGGCCAGCTGGCTGCAGATACAGAA 141
 QY 219 CysThrSerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGly 238
 DB 142 TGTACTTACGCTGTGATCTGTGAGCCTTAGAAATCATTTTCTGGAATGTTCCAGGA 201
 QY 239 MetLeuLeuHisThrValArgSerGlnGluTrpIleAlaAsnSerSerAlaIleIle 258
 DB 202 ATGAAACTGAAACATACATCAGATCTCAGAGATGAGAGCAAAAGTTCCTCTATAT 261
 QY 259 AspHisIlePheAlaSerLeuAlaValAlaAlaIleProAlaTYrHisLeuArg 278
 DB 262 GATCACATATTTGCCATGAAGCATGTGTGAATGCCCAATTCAGCCTTACCTTAGA 321
 QY 279 AspLeuIleLeuSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAla 298
 DB 322 GACCTTATCAAAAGCATGCTTCATGATGATCCAGAGAGAAATCTCTGTAATGGCA 381
 QY 299 LeuCysSerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeu 318
 DB 382 TTGTGAGCCCATTTCTTATGATCTCTTTGGCCCTCATATGGAATGATGTGATGCTT 441
 QY 319 ProThrProValLeuArgLeuLeuAlaValLeuAspAspTYrLeuGlnGluGlu 338
 DB 442 CCCATCTCAGTGTAAAGTCTGTAATGTGTGAGATGATATTAATTCAGAAATGAGAG 501

QY	339	GIUUTYRGUASBPVALVALGUAASYPVALLYSGUGUGUCYSGULNYSTYRGYPROVALVAL	358
QY	339	GIUUTYRGUASBPVALVALGUAASYPVALLYSGUGUGUCYSGULNYSTYRGYPROVALVAL	358
Db	502	GAATATGAAAGATGTTGTGAAAGATGTAAAGAGAGGTGTCAAAAATATGACCACTGTGTA	561
QY	359	SEIUDEUVALPOLYVSGUASBPVROGVALYRGYGVJVALPHEVALGIUTYRNAASN	378
Db	562	TCTCTACTGTTGTCCAAAGGAAATCTCTGGCAGAGCAAAAGTCTTGTGATGATCAAT	621
QY	379	ALAAGIAPSPSERLYSALALAGLNLVALLEUETHRGILYA-GMETPHEASPGLYLPSHE	398
Db	622	GCTGGTGATTCCAAAGCTGCACAGAAATTAAGTACTGACAGAGATGTTGATGGGAAGTTT	681
QY	399	VALVALAATATRPHELYPROLEUSERALATYLYSARCGLYTYRLEUTYRGINTHREU	418
Db	682	GTGTGGCTACATCTTACCCGTGAGTGCCTACCAAGAGGATCTGTATCAAACTTGG	741
QY	419	LEU 419	
Db	742	CTT 744	
RESULT 6			
LOCUS	CO771206	823 bp	linear
DEFINITION	testis EST02303 Testis cDNA Library Gallus gallus cdna 3', mRNA		
ACCESSION	CO771206		
VERSION	CO771206.1	GI:50972473	
KEYWORDS	EST,		
SOURCE	Gallus gallus (chicken)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.		
REFERENCE	1 (bases 1 to 823)		
AUTHORS	Shin,J.		
TITLE	Testis cDNA Library		
JOURNAL	Unpublished (2004)		
COMMENT	Contact: jihye Shin		
	Division of Animal Genetic Engineering		
	School of Agricultural Biotechnology, Seoul National University		
	San 56-1, Sillim-dong, Gwanak-gu, Seoul, 151-742, Korea		
	Tel: +82-31-479-7557		
	Fax: +82-31-479-7550		
	Email: paecljojh@hanmail.net		
	Seq primer: T3		
FEATURES			
source	Location/Qualifiers		
	1..823		
	/organism="Gallus gallus"		
	/mol_type="mRNA"		
	/strain="Korean Native chicken"		
	/db_xref="taxon:9031"		
	/sex="Male"		
	/tissue_type="Testis"		
	/dev_stage="26weeks"		
	/clone_lib="Testis cDNA Library"		
	/note="Vector: Uni-ZAP XR vector; Site_1: XhoI; Site_2: EcoRI; The libraries were synthesized using the ZAP(R)-cDNA synthesis method (Stratagene). cDNA was prepared, size-fractionated and inserted into the Uni-ZAP XR vector using XhoI linker-primer and EcoRI adaptor. After in vivo excision with E. coli strain SOLR, cDNA libraries from testis contain each insert size range of 0.5-3-kb (n=20), and 0.5-3.5kb (n=20)."		
ORIGIN			
Alignment Scores:			
Pred. No.:	3..61e-112	length:	823
Score:	1170..00	Matches:	229
Percent Similarity:	84.72%	Mismatches:	15
Best Local Similarity:	79.51%	Indels:	22
Query Match:	53.04%	Gaps:	2
DB:	7		

US-10-798-532-2 (1-419) x CO0771206 (1-823)
 Oy 47 GlySerProPro-----GlyAlaLeuLysGlnPheLeuProProGlyThrThrGly 63
 Db 6 GGAGCTCCACCGCGGTGGCGGCGCGCTCTGAACTAGTAGATCCCCCGGGCTGCAGAAAT 65
 Oy 64 ALaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeuGluGlnLeu 83
 Db 66 TCGGCACGAGCG----- 77
 Oy 84 GlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPheSerProAsn 103
 Db 78 -----GCCAACATCGTAGTACTGTGACGGCGTGTCCCAACCAACCACTACTCGGCCAAC 128
 Oy 104 ValProSerArgGlyLeuLeuLeuGlnLeuLeuLeuAspValSerValSerGluLeuLeuLeu 123
 Db 129 GCGCCGTCCTCCCGCTCCGTCTGTGAACTGTGTGAATATGACCGGTGTGAGCTGTCTGT 188
 Oy 124 TyrSerSerHisGlnGlyCysSerMetTyrMetIleGlnHisCysValAspAspValLeu 143
 Db 189 CACTTCGACCAACCGAGGCTGTCCATGTGATGATGCCAGCTATCGCCCGCGAGTCTGT 248
 Oy 144 GluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaPheLeuLysProArgAsn 163
 Db 249 GAAGCGCTGGCTTCTCTGCACCAACAAGGCTATCGTACACGACGACCTGAAGCCACGCAAC 308
 Oy 164 IleLeuTyrPheSerIleGluAsnGluCysPheLysLeuLeuIleAspPheGlyLeuSerPheLys 183
 Db 309 ATCTGTGGACCGCGCGAGAGAGATGTGTTAAAGCTCATGACTTTGACCTTAAGCTTCAAA 368
 Oy 184 GluGlyAsnGlnAspValLysTyrTleGlnThrAspGlyTyrArgAlaProGluAlaGlu 203
 Db 369 GAGGGGAAATCGAGATGTGAATATATTCACAAACAGACCGGATATCGGGCTCCAGAGCGAGA 428
 Oy 204 LeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThrSerAlaVal 223
 Db 429 CTGCAGAACTGCTTACGACAGGCGAGGCTCCAGAGTAGAGACGGAATGACTCTGTCTGTG 488
 Oy 224 AspLeuTyrPheSerLeuGlyIleIleLeuLeuGlnMetPheSerGlyMetLysValLeuHis 243
 Db 489 GATCTGTGAGATCTGGAGATGTTGTTTACTGAAATGTTCTCGAGGAAGAATCTGAACAT 548
 Oy 244 ThrValArgSerGlnGluTyrLysAlaAsnSerSerAlaIleIleAspHisIlePheAla 263
 Db 549 ACACTCCCAATTCACGAATGGAAGAACAACAGTTCTGCCATCATCGATTTTGTGCC 608
 Oy 264 SerLysAlaValAlaAsnAlaAlaIleProAlaTyrHisLeuArgAspLeuIleLysSer 283
 Db 609 AGTGAGAGGGGTGGTTAATTCAGCCATTCAGACTTATCACTCAAGACCTTATTAATAAGC 668
 Oy 284 MetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCysSerProPhe 303
 Db 669 ATGCTTCATTTGACCAAGGACCGAGGCTCTGTGAAAAGGTTTATGACGCCCATTC 728
 Oy 304 PheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThrProValLeu 323
 Db 729 TTCAGCAATTCCTTGTCTCCCAATTTGAAGATTGGTGAAGCTCCCAAGCCGTGTGCTG 788
 Oy 324 ArgLeuLeuAsnValLleuAspAsp 331
 Db 789 AGCGCTGCGAATGTTCTAAACGAT 812
 RESULT 7
 LOCUS BU447831 779 bp mRNA linear EST 29-NOV-2002
 DEFINITION 60376757.F1 CSEGRBN13 Gallus gallus cDNA clone CHESt695n17 5', mRNA
 sequence.
 ACCESSION BU447831
 VERSION BU447831.1 GI:25937142
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

1 (bases 1 to 779)
Boadman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
Ford, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
A Comprehensive Collection of Chicken CDNAS
Curr. Biol. 12 (22), 1965-1969 (2002)
2235534
12445392
Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES
source

1. 779
Location/Qualifiers
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Layer"
/db_xref="taxon:9031"
/clone="CHEST695n17"
/sex="Female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="CSEDBEN13"
/note="Organ: ovary; Vector: pBluescript II KS(+); Site_1:
ECORI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methyated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunt-ended, ligated to NotI adapters, digested with
ECORI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

ORIGIN

Alignment Scores:
Pred. No.: 2,31e-111 Length: 779
Score: 1162.00 Matches: 224
Percent Similarity: 93.02% Conservative: 16
Best Local Similarity: 86.82% Mismatches: 18
Query Match: 52.67% Indels: 1
DB: 5 Gaps: 0

US-10-798-532-2 (1-419) x BU447831 (1-779)

158 AapLeuLyProArGaEniLeuTPSeRaAGluuAnGluCyPhLeuLeuLeaAp 177
6 GACCTGAAGCAAGCAATCTCTGTGAGCGCGGAGAGAGAGCTTTAAGCTTATGAC 65
178 PheGluLeuSerPheLySgLuGluYaenGluAapValLySTyrlleGlnThraPglYTy 197
66 TTGGACTTGTAGCTTCAAGAGGGAATCAGAGATGTGAATAATATTCAAGACGGGTAT 125
198 ArgAlaProGluAlaGluLeuGluNaenCyLeuAlaGluAlaGluLeuGluNaenAap 217
126 CGGGCTCCAGAGCAAGCAAGTGTGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 185
218 GluCyThSerAlaValAapLeuTPSeRleuGlylleleuLeuGluGluMetPheSer 237
186 GAATGATACCTCTGTCTCA 245
QY 238 GtMetLyLeuLyShiThrValArgSerGlnGluTrpLyAlaAnSerSerAlaIle 257
|||||

DB 246 GGATGAAACTGAAACATACAGTCCATCTCAGAGATGAGAGACAAACAGTTCCCATC 305
QY 258 IleaPHisilePheAlaSerLySAlaValAlaAenAlaAlaIleProAlaTyrlHleu 277
DB 306 ATCGATCCGATTTTGGCCAGTGAAGGGGTGTATTTACAGCATTCAGCTTATACCTC 365
QY 278 ArgAapLeuLyLeuSerMetLeuHisAapAapProSerArgArgIleProAlaGluMet 297
DB 366 AGAGACCTTATTAAGACATGCTTCATTTGTGACCAAGCAGCGAGCTGTGCTGTAAG 425
QY 298 AlaLeuCySerProPhePheSerIleProPheAlaProHisIleGluAapLeuValMet 317
DB 426 GCTTATGAGCCCATCTTCTTCAGCATCTCTGCTCCCATATTAAGATTCGATGATG 485
QY 318 LeuProThProValLeuArgLeuLeuAenValLeuAapAapTyrlleuGluYaenGlu 337
DB 486 CTCGCCAGCCCTGTGTGCTGAGCTGTAATGTTCTTAAGGATGCTTCTGTGAGGTGA 545
QY 338 GluGluTyrgLuAapValAlaGluAapValLySgLuGluCyGlnLyTyrlGlyProVal 357
DB 546 GAAGATATGAGATATATCTCTGAGAGACATTAAGGAGAGAGTCAAGAAATATGACCGGTG 605
QY 358 ValSerLeuLeuValProLyGluAapProGluYaenGluAlaGluTyrlAla 377
DB 606 GTTCTCTGCTTATTCAGAGAGAAATCTGTAAAGCCAAAGCTTTGT -GAATATGCA 664
QY 378 AenAlaGluAapSerLySAlaAlaGluLySLeuLeuThrGluYArgMetPheAapLyLyS 397
DB 665 AATGCTGAGATTCCAAGACTGCCAAGAAATCTGACAGCTGAGCAAAATTTTATGAGCAAG 724
QY 398 PheValAlaIleThPheTyrlProLeuSerAlaTyrlYArgGlyTyrlleuTy 415
DB 725 TCTGCTGTGCTTACGATNTTACCACTGAGTGTCTATTAAGAGATATCTGTAC 778

RESULT 8
LOCUS COT73091 896 bp mRNA linear EST 04-AUG-2004
DEFINITION testis cDNA library Gallus gallus CDNA 3', mRNA
sequence.
ACCESSION COT73091
VERSION COT73091.1 GI:50974358
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 896)
Shin, J.
Testis cDNA library
Unpublished (2004)
Contact: JiHy Shin
Division of Animal Genetic Engineering
School of Agricultural Biotechnology, Seoul National University
San 56-1, Shillim-dong, Gwanak-gu, Seoul, 151-742, Korea
Tel: +82-31-479-7557
Fax: +82-31-479-7550
Email: jaeclioh@hannam1.net
Seq primer: T3

FEATURES
source

1. 896
Location/Qualifiers
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Korean Native chicken"
/db_xref="taxon:9031"
/sex="Male"
/tissue_type="Testis"
/dev_stage="26weeks"
/clone_lib="Testis cDNA library"
/note="Vector: Uni-ZAP XR vector; Site_1: XhoI; Site_2:
EcoRI; The libraries were synthesized using the
ZAP(R)-cDNA synthesis method (Stratagene); cDNA was
prepared, size-fractionated and inserted into the Uni-ZAP

XR vector using XhoI linker-primer and EcoRI adaptor.
After in vivo excision with E. coli strain SOIR, cDNA
libraries from testis contain each insert size range of
0.5-3kb (n=20), and 0.5-3.5kb (n=20)."

ORIGIN

Alignment Scores:

Pred. No.:	5 89e-111	Length:	896
Score:	1159.00	Matches:	240
Percent Similarity:	81.90%	Conservative:	18
Best Local Similarity:	76.19%	Mismatches:	30
Query Match:	52.54%	Indels:	27
DB:	7	Gaps:	2

US-10-798-532-2 (1-419) x C0773091 (1-896)

```

QY 46 ProglYser-PropPro-----GlyAlaLeuYsgInPheLeuProProglYThrTh 62
DB 11 CCGGNAACCTCCACCGCGTGGCGGCGCGCTTAAAGAACTAGATGCCCGGGCTGCAG 70
QY 62 rclYAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 82
DB 71 GAATTCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 86
QY 82 nleuGInglYHlsArgAsnIleValThrLeuYrGlYValPheThrIleHisPheSerPr 102
DB 87 -----CGCAACATCTGACCTCTGACCGCGCTTACCAACCACTACTCTGGC 133
QY 102 oAenValProSerArgCyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 122
DB 134 CAACGCGCGCGTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 193
QY 122 uLeuYrSerSerHisGInglYCySerMetTPrMetIleGInHisCyAlaArgAspVa 142
DB 194 GCGCACTCCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 253
QY 142 lLeuGInglAlaLeuAlaPheLeuHisIleGInglYrValHisAlaAspLeuYrProAr 162
DB 254 GCTGGAGAGCCCTGCTCTCTGACCAACAGGCTACGTCAGCAGACCTTAAGCCACG 313
QY 162 gAenIleLeuYrSerArgIleGInglYCyPheLeuLeuLeuLeuLeuLeuLeuLeuLeu 182
DB 314 CAACATCTGCTGAGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 373
QY 182 elYsgInglYArgGInAspValYsYrIleGInThrAspGlyYrArgAlaProGluAl 202
DB 374 CAAGAAGGAGAAACAGAGATGTGAATATATATCAACAGACGGGTATCGGGCTCC 433
QY 202 aGluLeuGInAsnCyLeuAlaGInAlaGInleuGInSerAspThrGluCySerHisSer 222
DB 434 AGAATCTGCAAGATCTGCTACAGAGGAGGCTCCAGAGTGAAGAGAGAGAGAGAGAG 493
QY 222 aValAspLeuYrSerSerleuGlyIleIleLeuLeuGInMetPheSerGlyMetIleYs 242
DB 494 TGTGATCTGTGAGCTGTGAGATGTGTTTACTGGAATGTTCTCAGGAATGAAGTGA 553
QY 242 sHisThrValArgSerGInglYrTrpYsAlaAsnSerSerAlaIleIleAspHisIlePh 262
DB 554 ACATTAAGTCCATCTCCAGAGATGAAGACAAACATTTCTGCATATCATGATGCATTTT 613
QY 262 eAlaSer-LysAlaValAlaAsnAlaAlaIleProAlaYrHisLeuArgAspLeuIle 282
DB 614 TGGCAATGAAGGGGGGTGTTTATTCAGCCATTCAGCTTATTCAGAGACCTTATTA 673
QY 282 ySerMetLeuHisAspAspProSerArgIleProAlaGInMetAlaLeuCySerP 302
DB 674 AAAGCATGCTTCACTTGACCAAGCAAGGAGGAGCTCTGCTGAAGAGGCTTATGACGC 733
QY 302 rOpheHisSerIleProPheAlaProHisIleGInAspLeuValMetLeuProThrProY 322
DB 734 CATTTCTGAGCATCTCCCTTCTCCCATATGAAAGATTGAGATGCTCCCAAGCCCTG 793
QY 322 alLeuYrGluLeu-LeuAsnValLeuAspAspTyrlLeuGlyAsn-GInGInglYrGl 341

```

```

DB 794 TCCTAGAGCTGACTGAATGTTTAAAGCATGCTTCTTCGACAGTGAAGAGATACGA 853
QY 341 uAspValValGlu-AspValYsGInglYCyGln 352
DB 854 AGAATCCCTGAGACATAGAGGAGAGAGATGTCAG 888

```

RESULT 9

CN232453

LOCUS CN232453 791 bp mRNA linear EST 09-APR-2004

DEFINITION WLB074D11.ab1 Wtrestis Gallus gallus cDNA 5', mRNA sequence.

ACCESSION CN232453

VERSION CN232453.1 GI:46336197

KEYWORDS EST

SOURCE Gallus gallus (chicken)

ORGANISM Gallus gallus

REFERENCE Savolainen, P., Fitzsimmons, C.J., Arvestad, L., Andersson, L. and

AUTHORS Savolainen, P.

TITLE EST analysis of brain and testis cDNA libraries from White Leghorn

COMMENT Unpublished (2004)

DEPARTMENT OF Biotechnology

ROYAL INSTITUTE OF TECHNOLOGY, KTH

SE-106 91 Stockholm, SWEDEN

TEL: +46 (0)8 5537 8481

FAX: +46 (0)8 5537 8335

EMAIL: Peter.Savolainen@biotech.kth.se

Seq primer: M13 reverse primer.

Location/Qualifiers

1..781

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="White Leghorn"

/db_xref="taxon:9031"

/sex="male"

/lab_host="ElectroMAX DH10B (Invitrogen)"

/clone_lib="Wtrestis"

/note="Organ: testis; Vector: pSPORT-1; Site 1: Hind III;

Site 2: EcoRI; The cDNA libraries were created with the

Superscript Plasmid System (Invitrogen)."

ORIGIN

Alignment Scores:

Pred. No.:	2.61e-108	Length:	781
Score:	1133.00 <td>Matches:</td> <td>225</td>	Matches:	225
Percent Similarity:	93.10% <td>Conservative:</td> <td>18</td>	Conservative:	18
Best Local Similarity:	86.21% <td>Mismatches:</td> <td>18</td>	Mismatches:	18
Query Match:	51.36% <td>Indels:</td> <td>3</td>	Indels:	3
DB:	7	Gaps:	0

US-10-798-532-2 (1-419) x CN232453 (1-781)

```

QY 109 lLeuLeuLeuGInLeuLeuAspValSerValSerGluLeuLeuYrSerSerHisGIn 128
DB 2 CTGCTGCTGAGACTGCTGATATCAAGCGTGTGAGCTGCTGCTCACTCCAGAACCG 61
QY 129 GlYCySerMetTPrMetIleGInHisCyAlaArgAspValLeuGluAlaLeuAlaPhe 148
DB 62 GGCTGCTCATGTGATGATGATCCAGCACTGCGCGCGAGAGTCTTGAAGCCCTGACCTT 121
QY 149 lLeuHisIleGInglYrValHisAlaAspLeuYsProArgAsnIleLeuYrSerAla 168
DB 122 CTGCAACCAAGAGCTACGACCAAGCCTGAAGCCAGAACTCTGTGAGCGCG 181
QY 169 GluAsnGluCyPheLeuYsLeuIleAspPheGlyLeuSerPheYsGInglYAsnGlnAsp 188
DB 182 GAGAGAGAGTGTCTTAAAGCTCATGATGATTAAGCTTCAAAAGAGGGAGATCAGAT 241

```

```

QY 189 ValleyTyrIleGlnThrAspGlyTyrArgAlaProGluAlaGluLeuGlnAenCyLeu 208
Db 242 GTGAAATATATTCACACAGACGGGTATCGGGCTCCAGAGCAGAACTGCAGACTGCTTA 301
QY 209 AlAGlnAlaGlyLeuGlnSerAspThrGluCysThrSerIleValAspLeuTppSerLeu 228
Db 302 GCACAGGACGAGGCTCCAGAGTGAACGAAATGACCTGCTGATGATGATGATGATGATG 361
QY 229 GlyIleLeuLeuGlnMetCpnsSerGlyMetLeuLeuYshIsthrValAspSerGln 248
Db 362 GGAATCTTTTATTCGAAATGCTTCTCAGAAATGAACTGAACTGAACTGAACTGAACTG 421
QY 249 GluTyrPylalaenSerSerAlaIleAspHisIlePheAlaSerIleValVal 268
Db 422 GAATGAGACAAACAGTCTGCTCATCATCATCATCATCATCATCATCATCATCATCATCAT 481
QY 269 AenAlaAlaIleProAlaTyrHisLeuArgAspLeuIleYserMetLeuHisAspAsp 288
Db 482 AATTCAGCCATTCAGTTCATCCTCAGAGACCTTATTAAGCATGCTTCATGTCGAC 541
QY 289 ProSerArgArgIleProAlaGluMetAlaLeuCysSerProPhePheSerIleProPhe 308
Db 542 CAGGCAAGACAGCCCTCTGCTGAAAGCTTATGACGCCCATTCCTTCAGCATTCCTTT 601
QY 309 AlaProHisIleGluAspLeuValMetLeuProThrProValLeuArgLeuLeuAenVal 328
Db 602 GCTCCCATATTCAGATTCGATGATGATGATGATGATGATGATGATGATGATGATGATG 661
QY 329 LeuAspAspAspPyrLeuGluYasnGluGluGluGluGluGluGluGluGluGluGlu 348
Db 662 CTAGGCAATCTCTCTGCTGAGAGAGAGA-GAAATACAAAGATATCTGGAAGA-ATPAGG 719
QY 349 GluGluCysGlnIleYsrGlyProValValSerLeuValProYsgIleAenProGly 368
Db 720 GAGAGGTGTAGAAATATGAGCCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 778
QY 369 Arg 369
Db 779 AAA 781

RESULT 10
BUS41306 932 bp mRNA linear EST 13-SEP-2002
LOCUS AGENCOURT_10327419 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6572970
DEFINITION BUS41306
5', mRNA sequence.
ACCESSION BUS41306.1 GI:22851747
VERSION
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 932)
NIH-MGC http://mgc.ncl.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LHC8765 row: 0 column: 18
High quality sequence stop: 631.
Location/Qualifiers
1..932
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6572970"

```

```

/cisue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH MGC 40"
/notes="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GSCACGAG(6). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

```

```

ORIGIN
Alignment Scores:
Pred. No.: 8,94e-106 Length: 932
Percent: 110.00 Matches: 228
Score Similarity: 87.22% Conservative: 4
Best Local Similarity: 85.71% Mismatches: 18
Query Match: 50.32% Indels: 16
DB: Gaps: 2

```

US-10-798-532-2 (1-419) x BUS41306 (1-932)

```

QY 1 MetAlaGlySerGlyCysAlaIleTrrGlyAlaGluProProArgPheLeuGluAlaPheGly 20
Db 149 ATGGCGGATCCCGCTGCGCGCGGAGCGCGCGCGGCTTCTGAGGCGCTTCGGG 208
QY 21 ArgLeuTrrGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
Db 209 CGGCTGTGCGAGTACAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 268
QY 41 ArgCysArgIleYasnProGlySerProProGlyAlaLeuYsgIlePheLeuProProGly 60
Db 269 CGCTGTGCGGCAACCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 328
QY 61 ThrThrGlyAlaAlaAspSerAlaIleGluTyrGlyPheArgGlyGluAlaAlaLeu 80
Db 329 ACCACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 388
QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
Db 389 GAAAGCTTCAGAGGTCACAGAAACATCGATTTGTATGAGAGTGTATCAATCCACTTT 448
QY 101 SerProAenValProSerArgCysLeuLeuGluLeuLeuAspValSerValSerGlu 120
Db 449 TCTCCAAATGTCCATCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 508
QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrrMetIleGlnHisCysAlaArg 140
Db 509 TTGCTCTTATTCACGTCACAGCGCTGTTCATGTGATATACAGCATGTGCGCGA 568
QY 141 AspValLeuGluAlaLeuAlaPheLeuHisIsgIleGlyTyrValHisAlaAspLeu-Ly 160
Db 569 GATGTTTGGAGCGCTTCTTCTTCATCATGAGGCTATGTCATGCGGACCTCCCA 628
QY 160 sProArgAsnIleLeuTrrSerAlaGluAsnGluCysPheYsLeuIleAspPheGlyLe 180
Db 629 ACCACGTAATATGAGTGTGAGTGCACAAAGATGAATGTTTAAATCATGACTTGGACT 688
QY 180 u-SerPheYsgIleGluYasnGlnAspValYsTyrIleGlnThrAspGlyTyrArgAlaP 200
Db 689 TTAGCTTCAAAAGAACATCGAGATGTAATATTCACAGACACGCGGATTCGCGCTC 748
QY 200 roGlu-AlaGluLeuGlnAsnCysLeuAlaGlnAla-GlyLeuGlnSerAspThrGlyCly 219
Db 749 CAGAAAACGAATTCGAAATGCTTGGCCCAACGCTTGCGCTCGCAAGATACAAATG 808
QY 219 sThrSer-AlaValAsp-LeuTrrSer-----LeuGlyIleIleLeuLeuGluMetPhe 236
Db 809 TACCTTCAGCTGTGATCTGTGAGAGCCCTAAGAAATCATTTTACGAAATATGATGTC 868
QY 237 SerGlyMetLeuYsLeuYshIsthrValAspSerGlnGluTrrPylalaenSerSerAla 256

```

Db 869 CAGGATG-----GAAACTGGAAACCTACGCGTCCAGA 904

QY 257 llelle 258

Db 905 ATCTTC 910

RESULT 11

BU385761

LOCUS 717 bp mRNA linear EST 28-NOV-2002

DEFINITION 60382650F1 CSECHN75 Gallus gallus cDNA clone CHEST533d13 5', mRNA

ACCESSION BU385761

VERSION BU385761.1 GI:25893762

KEYWORDS EST.

SOURCE Gallus gallus (chicken)

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

REFERENCE 1 (bases 1 to 717)

AUTHORS Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E., Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.

TITLE A Comprehensive Collection of Chicken CDNAS

JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)

MEDLINE 2235534

PUBMED 12445392

COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612060930
Fax: 01612260409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES

source

1.717

Location/Qualifiers

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="white Leghorn, H1sex"

/db_xref="taxon:9031"

/clone="CHEST533d13"

/dev_stage="36"

/lab_host="DH10B"

/clone_1b="CSECHN75"

/notes="Organ: trunks; Vector: pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunted, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN

Alignment Scores:

Pred. No.: 1.59e-105 length: 717

Score: 1106.00 Matches: 210

Percent Similarity: 93.72% Conservative: 14

Best Local Similarity: 87.87% Mismatches: 14

Query Match: 50.14% Indels: 1

DB: 5 Gaps: 0

US-10-798-532-2 (1-419) x BU385761 (1-717)

QY 134 Met15Gln153Cys15Arg25Pval15Leu15Ala15Leu15Ala15Phe15Leu15His15Gln15Gly 153

Db 3 ATATCATCAGACATCGCGCCGCGAAGCTCTTGGAAGCCCTTCCCTGACCAACAAAGCC 62

QY 154 TyrValHisAlaAspLeuYsProArgAsnIleLeuTrpSerAlaGluAsnGluCysPhe 173

Db 63 TACGTGACGACGACCTGTAAGCAAGCAATCCTGTGAGCGCGAGAGAGCTTT 122

QY 174 IysLeuIleAspPheGlyLeuSerPheIysGluGlyAsnGlnAspValIysTyrIleGln 193

Db 123 AAGCTCATGACTTGTGACTTGTGCTTCAAGAGGGGAAATCAGAGATGTAAATATATTCCA 182

QY 194 ThrAspGlyTyrTyrAlaPArgGluAlaGluLeuGlnAsnGlyLeuAlaGlyLeu 213

Db 183 ACAGAGCGGTATCGGCTCCACAGAGCAAGCACTGCACTGACAGAGCGGCTC 242

QY 214 GlnSerAspThrGluCysThrSerAlaValAspLeuTrpSerLeuGlyIleLeuLeu 233

Db 243 CAGAGTACAGCGAATGTAATCTCTCTGTGATCTGTGAGCTGGAATCTTTACTG 302

QY 234 GluMetPheSerGlyMetIysLeuIysHisIleThrValArgSerGlnIleTrpIysAlaAsn 253

Db 303 GAAATGTTCTCAGAATGAATCACTAAGCAATACAGTCAATCTCAGAAATGMAAGCAAAAC 362

QY 254 SerSerAlaIleIleAspHisIlePheAlaSerIysAlaValAlaAsnAlaIlePro 273

Db 363 AGTTCTGCATCATCATGATCGCATTTTTCAGATGAAGGGGTGTAAATTCACGCTTCCA 422

QY 274 AlaTyrHisIleLeuArgAspLeuIleIysSerMetLeuHisAspAspProSerArgIle 293

Db 423 GCTTATACCTCAGAGACCTTATTAAGCATGCTTCTTGTGACCAAGCAAGCAAGCC 482

QY 294 ProAlaGluMetAlaLeuCysSerProPhePheSerIleProPheAlaProHisIleGlu 313

Db 483 TCTGCTGAAAGAGCTTTATGAGCCCATCTTCACATTCCTTGTCTCCCATATTGAA 542

QY 314 AspLeuValMetLeuProThrProValLeuArgLeuLeuAsnValLeuAspAspTyr 333

Db 543 GATTGGTGATGCTCCACGCGCTGTGCTGAGCTGCAATGTTCTAAGCATCTTCT 602

QY 334 LeuGlyAsnGluGluGlyTyrGluAspValIleGluAspValIleGluGluCysGlnIys 353

Db 603 CTGCAAGAGTGAAGAAATACCAAGATATCTCTGAAAGCATTAAGGAGAGTGTGAAA 662

QY 354 TyrGlyProValIleSerLeuLeuValProIysGluAsnProGlyIysGlyGln 371

Db 663 TATGACCGGTGTTTCTTCTTATTCGAAAGCAATCTGTAAAGGCCAA 717

RESULT 12

BE895119 899 bp mRNA linear EST 20-OCT-2000

LOCUS 601436041F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3920939 5',

DEFINITION mRNA sequence.

ACCESSION BE895119

VERSION BE895119.1 GI:10358193

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

AUTHORS NIH-MGC http://mgi.mc.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: ATCC/DC/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at:
http://image.llnl.gov
Plate: ILMA9753 row: b column: 12
High quality sequence stop: 626.

FEATURES

Location/Qualifiers


```

source
1. .899
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3920939"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_11b="NIH_MGC_72"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally. primer: Oligo dT. Average insert size 2 kb. Library constructed by life technologies."

ALIGNMENT SCORES:
Pred. No.: 8 45e-104 Length: 899
Score: 1091.00 Matches: 210
Percent Similarity: 99.53% Conservative: 0
Best Local Similarity: 99.53% Mismatches: 0
Query Match: 49.46% Indels: 1
DB: 2 Gaps: 0

US-10-798-532-2 (1-419) x BE895119 (1-899)

QY 113 LeuLeuAspValSerValSerGluLeuLeuTySerSerHisGlnGlyCysSerMet 132
DB 1 CTCCTGGATGTCAGTGTTCGGAATGCTCTTATATCCAGTCCAGCAGGGTGTCCATG 60

QY 133 TrpMetIleGlnHisCysAlaArgAspValLeuGluAlaLeuAlaPheLeuHisGln 152
DB 61 TGGATATATACGACATGTGTCGCGAGATGTTTGAGGCCCTTCTTCTTCATCATGAG 120

QY 153 GlyTyValHisAlaAspLeuLeuProArgAsnIleLeuTrpSerAlaGluAsnGlyCys 172
DB 121 GGGTATGTCATCGGAGCCTCAACACGATCATATGTGCGATGCGAGAGATGATGT 180

QY 173 PheLeuLeuIleAspPheGlyLeuSerPheLeuGluGlyAsnGlnAspValIleTyTyr 192
DB 181 TTAAATCTCATGTGACTTGTGAGCTTAGCTTCAAGAAAGGCAATCAGATGTAAAGTATAT 240

QY 193 GlnThrAspGlyTyTrpAlaProGluAlaGluLeuGlnAsnGlyLeuAlaGly 212
DB 241 CAGACAGACGGGATATGCGGCTCCAGAAAGCAAAATGCAAAATGCTTGCGCCAGCGTGGC 300

QY 213 LeuGlnSerAspThrGluCysTrpSerAlaValAspLeuTrpSerLeuGlyIleLeu 232
DB 301 CTCGACAGATGATACAGAAATGTACTCTGAGTGTGCGAGCCTAGGAATCATTTTA 360

QY 233 LeuGluMetPheSerGlyMetLeuLeuValHisThrValArgSerGlnGluTyrPheAla 252
DB 361 CTGGAAATGTCTTCAGGAATGAAATGAAACATACAGTCTCAGATCTCAGGAATGAAAGGCA 420

QY 253 AsnSerSerAlaIleLeuAspHisIlePheAlaSerTyValValAlaAlaAlaIle 272
DB 421 AACAGTTCGCTATATATGATCAATATTTGCCAGTAAGACAGTGGTAATGCGCAATT 480

QY 273 ProAlaTyHisLeuAlaAspLeuLeuLeuSerMetLeuHisAspAspProSerArgArg 292
DB 481 CCAAGCCTATCACTTAAGAGACCTTATCAAAAGCATGCTTATGATATCCAAACAGAGA 540

QY 293 IleProAlaGluMetAlaLeuCysSerProPhePheSerIleProPheAlaProHisIle 312
DB 541 ATTCCTGCTGCTAAATGGCATTTGTGACAGCCCATCTTTAGCATTCCTTTTGCCCTCATATT 600

QY 313 GluAspLeuValMetLeu-ProThrProVal 322
DB 601 GAAAGATCTGATCATGCTTTTCCACTCCAGTT 631

RESULT 13
BI562420 772 bp mRNA linear EST 05-SEP-2001
LOCUS BI562420
DEFINITION 60326316F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5298674 5',
mRNA sequence.

```

```

ACCESSION BI562420
VERSION BI562420.1 GI:15449747
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 772)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@db-rc@mail.nih.gov
Tissue Procurement: Miklos Palokvits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shihaki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL1755 row: P column: 03
High quality sequence stop: 727.
Location/Qualifiers
1. 772
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5298674"
/lab_host="DH10B"
/clone_11b="NIH_MGC_97"
/notes="Organ: testis; Vector: pBluescriptPR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
(gcgag); Oligo-dT primed using primer
5'-TTTTTTTTTTTNN-3', size-selected for average
insert size 2.2 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC library."

ALIGNMENT SCORES:
Pred. No.: 1.23e-102 Length: 772
Score: 1079.00 Matches: 212
Percent Similarity: 96.82% Conservative: 1
Best Local Similarity: 96.36% Mismatches: 5
Query Match: 48.91% Indels: 3
DB: 4 Gaps: 1

US-10-798-532-2 (1-419) x BI562420 (1-772)

QY 92 LeuTyGlyValPheThrIleHisPheSerProAsnValProSerArgCysLeuLeu 111
DB 108 TTGTATGAGGTGTTTCAATCCACTTTCTCCAAATGTCACACGCTGTGTTGCTT 167

QY 112 GluLeuLeuAspValSerValSerGluLeuLeuTySerSerHisGlnGlyCysSer 131
DB 168 GAATCTCTGATGTCACTGTTTCGGAATGCTCTTATATTCAGTACCAAGGTTGTTCC 227

QY 132 MetTrpMetIleGlnHisCysAlaArgAspValLeuGluAlaLeuAlaPheLeuHis 151
DB 228 ATGTGATGATACAGCATTTGTGCGAGATGTTTGAGAGCCCTTCTTCTTCATCAT 287

QY 152 GlyTyTyValHisAlaAspLeuLeuProArgAsnIleLeuTrpSerAlaGluAsnGly 171
DB 288 GAGGCGTATGTCATGCGGACCTCAACACGATACATATTTGTGAGTCAAGAAATGAA 347

QY 172 CysPheLeuLeuIleAspPheGlyLeuSerPheLeuGluGlyAsnGlnAspValIleTyTyr 191
DB 348 TGTTTAACTCATTTGACTTTGGACTTACCTTCAAAAGAGGCAATCAGGATGTAAAGTAT 407

```


QY 192 IleglnThrAspGlyTyrArgAlaProGluAlaGluLeuGlnAsnCySeuLeuAlaGlnAla 211
Db 408 ATTCAGACGACGCGGATCCGGCTCCAGAGCAGAAATGCAAAATTCCTGGCCCGCGCT 467
QY 212 GlyLeuGlnSerAspThrGluCysThrSerAlaValAspLeuTrpSerLeuGlyIleIle 231
Db 468 GGCTTCAGAGTATCAGAAATGATTAATCTGAGCTGTGATCTGAGAGCTTGAAGATCAT 527
QY 232 LeuLeuGlnMetPheSerGlyMetLeuValLeuHisThrValArgSerGlnGluTrpLys 251
Db 528 TTAATGAAATGTTCTCCAGAAATGAACTGAACATACATCTGAGATCTCCAGAAATGAG 587
QY 252 AlaAsnSerSerAlaIleIleAspHisIle-PheAlaSerGlyAlaValAlaAsnAlaI 271
Db 588 GCAAACAGTCTCTCTATTATGATCACTATTGTCAGTAAGAGCAGTGTGATGCGC 647
QY 271 alleProAlaTyrHisLeuArgAspLeuIleLysSerMetLeuHisAspAspSerAr 291
Db 648 AATTCAGCCTTACCTAAGAGCCTTATCAAAAGCATGCTTCATGATGATCCAGACAG 707
QY 291 gaArgIleProAlaGluMetAlaLeuCysSerProPhePheSerIleProPheAlaPro 310
Db 708 AAGAAATTCCTGCTGAATGGCATTTGGCAG-CCATCTTGAACA---TCCTTGCCCT 761
RESULT 14
BU205357 794 bp mRNA linear EST 25-NOV-2002
LOCUS 603104569F1 CSBQCHN03 Gallus gallus cDNA clone CHEST4118 5', mRNA
DEFINITION
sequence.
ACCESSION BU205357
VERSION BU205357.1 GI:25370848
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 794)
Boardman, P.E., Sanz-Exquerro, J., Overton, I.M., Burt, D.W., Boech, E.,
Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.,
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
2235534
JOURNAL MEDLINE
PUBMED 12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1. 794
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="CHEST4118"
/issue_type="whole embryo"
/dev stage="20-21"
/lab host="DH108"
/clone_1lb="CSBQCHN03"
/note="Organ: whole embryo; Vector: pBluescript II KS(+);
Site 1: EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunt-ended, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2

rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

ALIGNMENT SCORES:

Alignment Scores:	3.38e-102	Length:	794
Score:	1075.00	Matches:	221
Percent Similarity:	89.10%	Conservative:	16
Best Local Similarity:	83.08%	Mismatches:	27
Query Match:	48.73%	Indels:	4
DB:	5	Gaps:	0

US-10-798-532-2 (1-419) x BU205357 (1-794)

QY 82 GlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPheSer 101
Db 1 CAGCTCCGGGGGACCGCAACATCGTGACTGTACGGCGGTTCACCAACACATCACTCG 60
QY 102 ProAsnValProSerArgCysLeuLeuLeuGlnLeuLeuAspValSerValSerGluLeu 121
Db 61 GCCAACGGCCCGTCCCGCTGCTGCTGAGCTGTGGATATGACGCGTCTGAGCTG 120
QY 122 LeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArgAsp 141
Db 121 CTGCTGCACTCCAGCAACAGGCTCTCTCATGTGATATATCAGACTGCGCCGCGAC 180
QY 142 ValLeuGlnAlaLeuAlaPheLeuHisHisGlnGlyTyrValHisAlaAspLeuLysPro 161
Db 181 GTCTCGAAGACCTGCGCTTCCCTGACACCAAAAGGCTACGTGACACGACGCAAGCA 240
QY 162 ArgAsnIleLeuTrpSerAlaGluLysGluCysPheValLeuIleAspPheGlyLeuSer 181
Db 241 CGCAACATCTGCTGAGCGCGGAGGAGGAGCTTTAACTCATGACTTGGACTTACG 300
QY 182 Phe-LysGlnGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaProG 201
Db 301 TTCAAAGAGGGGAAATCAGATGTGAAATATATTCAAACAGCGGTATCGGGCTCCAA 360
QY 201 uAlaGluLeuGlnAsnCySeuLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThrSe 221
Db 361 GGCAGAACTGCAGAAATGCTGAGCAGGCGGCTCCAGATGAGACGGAATGTACTCTC 420
QY 221 rAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnMetPheSerGlyMetLysSe 241
Db 421 TGCTGTGATCTGTGAGTCTGGGAATGTTTAACTGGAATGTTCTCGAGATGAAGAACT 480
QY 241 uLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleAspHisIle 261
Db 481 GAAACATACAGTTCATCTCAGGAATGAAACAAACAGTTCTGCCATATCATGATCGCAT 540
QY 261 epPheAlaSerIlyAlaValAlaAsnAlaIleProAlaTyrHisLeuArgAspLeuI 281
Db 541 TTTTGCAGAGAAAGGGGTGTTAATCAGACCAATTCAGCTTATCACTCAGAGACCTTAT 600
QY 281 eLysSerMetLeuHisAspAsp-ProSerArgArgIleProAlaGluMetAlaLeuCys 301
Db 601 TAAACCATGCTTCATTTGACCAAGCAAAAGCAACCTCTGTGAAAAGCTTAATGCA 660
QY 301 exProPhePheSerIleProPheAlaProHisIleGlnAspLeuValMetLeuProThrP 321
Db 661 GCCCATTTTCAGCATTTCCCTTGCTCC-CAATTTAAATTTGGATGCTCCCAAG-C 718
QY 321 roValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlnGlnGlnGlnGlnGln 341
Db 719 CTGTGTGAGGCTGTGAAATGTTCTTAAGCGCAAGGTTCTCGAAGGTGAAGAAATACG 778
QY 341 LuAspValValGlu 345
Db 779 AAGAAATCTGGGAA 792
RESULT 15

CF593363
 LOCUS 834 bp mRNA linear EST 26-SEP-2003
 DEFINITION AGENCOURT 15622279 NIH_MGC.147 Homo sapiens cDNA clone
 IMAGE:30531268 5', mRNA sequence.
 ACCESSION CF593363
 VERSION CF593363.1 GI:36346826
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 834)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: CGAPbs-r@mail.nih.gov
 Tissue Procurement: Dr. Stefan Hansson
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
 and advice from Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: NDAB621 row: m column: 05
 High quality sequence stop: 660.
 Location/Qualifiers
 1. 834
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30531268"
 /tissue_type="Human Placenta"
 /lab_host="DH10B Tona"
 /clone_1lb="NIH_MGC_147"
 /note="Organ: Placenta; Vector: pBluescriptR; Site 1:
 all-XhoI; Site 2: BamH; Oligo-dT primed using primer
 5'-TTTTTTTTTTTTTTVN-3', size-selected for average
 insert size 2.3 kb and normalized to R0T 5. This is a
 primary library enriched for full-length clones and
 constructed using the Cap-trapper method (Carninci, in
 preparation). Library constructed by M. Brownstein
 (NIH/NHGRI, National Institutes of Health). Note: This is
 a NIH_MGC library."

ORIGIN

Alignment Scores:

Pred. No.:	5,24e-102	Length:	834
Score:	1073.50	Matches:	214
Percent Similarity:	95.13%	Conservative:	1
Best Local Similarity:	94.69%	Mismatches:	5
Query Match:	48.66%	Indels:	6
DB:	7	Gaps:	1

US-10-798-532-2 (1-419) x CF593363 (1-834)

QY 1 MetAAGlySerGlyCyValATTPGlyAlaGluProBArgPheLeuGluAlaPheGly 20
 Db 155 ATGGCGGGAATCCGGCTGCGCTGGGGCGCGAGCGCGGCTTTCTGAGAGCCTTCGGG 214
 QY 21 ArgLeuTrpGluValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
 Db 215 CGGCTGTGGCAAGGACGACCGGCTGTGGGTAGCGGCTCTCCGCTCGGTATCGGGTT 274
 QY 41 ArgCySerGlyAlaSerProGlySerProProGlyValAlaLeuTyrGlnPheLeuProProGly 60
 Db 275 CGCTGTGGCGCAACCTGTGCTGCGCCCGCGGCCCTCAAGCAATTCTTGTGGCGCAAGA 334
 QY 61 ThrThrGlyAlaAlaSerAlaAlaGluTyrGlyPheArgGlySerGluArgAlaAlaLeu 80

Db 335 ACCACCGGGGCTGCGGCTCTGCGCGCGATGATGTTCCGCAAGAGAGGGGCGGCTG 394
 QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
 Db 395 GAACAGTGTGAGGATCACAGAAACATCGTACTTGTATGAGTGTATTCACATCCACTTT 454
 QY 101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
 Db 455 TCTCCAAAGTGCATCACGCTGTCTGTGCTTGAACCTCTGGATGTCAGTGTTCGGA 514
 QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysValAlaArg 140
 Db 515 TTGCTCTTATTTCCAGTCACCGAGGTTGTTCATGTGATGATACAGCAATTGTGCCA 574
 QY 141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
 Db 575 GATGTTTGGAGGCTCTGCTTTCTTCATCATGAGGCTATGTCCATGCGAGCTCANA 634
 QY 161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGly-Le 180
 Db 635 CCACTTACATATTTGTGAGTGCAGAGAAATGATTTTAACTCATTTGACTTTGGA 694
 QY 180 uSerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArg-AlaP 200
 Db 695 AGGCTTCAAGAAAGAACATCAGGATGTAAAGTATTCAGACAGACGAGGTATCGGGCTC 754
 QY 200 roGluAlaGluLeuGln-AAsnCysLeuAla---GlnAlaGly-LeuGlnSerAspThrG 218
 Db 755 CAGAACACAGATTGCAAAAATTGCTTGGGCCCAAGGCTTGCCCTTCAGAGTATACAG 814
 QY 218 u-CysThr 220
 Db 815 AATGTACC 822

Search completed: November 29, 2004, 13:27.43
 Job time : 3172.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2004, 11:26:47 ; Search time 26 Seconds
(without alignments)
1550.570 Million cell updates/sec

Title: US-10-798-532-2

Perfect score: 2206

Sequence: 1 MAGSCAMGAEPFRFLBARG.....VATFYPPLSAVKKGYLQTL 419

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1052	47.7	200	2	148615
2	248.5	11.3	948	2	T24445
3	243.5	11.0	720	2	A96807
4	243.5	11.0	826	2	JC5153
5	241.5	10.9	387	2	T37758
6	240	10.9	388	2	S48879
7	235.5	10.7	367	2	S68680
8	234.5	10.6	354	2	T45138
9	234.5	10.6	433	2	S37790
10	234	10.6	435	2	S64950
11	233	10.6	298	1	A44878
12	233	10.6	298	2	A41227
13	233	10.6	369	2	A56492
14	230	10.4	373	2	T13024
15	230	10.4	376	2	S40471
16	230	10.4	427	2	JC5693
17	228	10.3	423	2	S43968
18	227.5	10.3	469	1	S17656
19	227	10.3	326	2	S23387
20	227	10.3	376	2	S40470
21	227	10.3	423	2	S43967
22	227	10.3	424	2	S71102
23	227	10.3	529	1	S71774
24	226.5	10.3	367	2	JC5252
25	226	10.2	346	1	I48157
26	226	10.2	406	1	KIHUCT
27	225.5	10.2	301	1	S19209
28	225.5	10.2	371	2	T14915
29	225.5	10.2	416	2	A48249

30	224.5	10.2	301	1	A48041	protein kinase (EC
31	224.5	10.2	678	2	T43539	spindle checkpoint
32	224	10.2	358	2	S23383	protein kinase (EC
33	223.5	10.1	335	2	T23050	hypothetical prote
34	223.5	10.1	342	2	T21098	protein kinase (EC
35	223.5	10.1	426	2	S43969	p54-beta stress-ac
36	223.5	10.1	464	2	S71104	protein kinase JNK
37	223.5	10.1	602	2	S60052	calcium-dependent
38	223	10.1	427	2	JC5694	stress-activated p
39	223	10.1	575	2	JC7794	lammer kinase homo
40	223	10.1	690	2	T38052	probable protein k
41	222.5	10.1	393	2	S51321	mitogen-activated
42	222.5	10.1	424	2	A55480	c-Jun amino-termi
43	222.5	10.1	473	1	A53036	Ca2+/calmodulin-de
44	222	10.1	346	1	I78840	protein kinase (EC
45	222	10.1	371	2	S60121	mitogen-activated

ALIGNMENTS

```

RESULT 1
148615
gene KIS protein - mouse (fragment)
C/Species: Mus musculus (house mouse)
C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C/Accession: I48615
R/Maucuer, A.; Camonis, J.H.; Sobel, A.
Proc. Natl. Acad. Sci. U.S.A. 92, 3100-3104, 1995
A/Title: Statmin interaction with a putative kinase and coiled-coil-forming protein dome
A/Reference number: I48282; MUID:95241452; PMID:7724523
A/Accession: I48615
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-200 <RES>
A/Cross-references: UNIPROT:P97343; EMBL:X82320; NID:G791078; PIDN:CAA57763.1; PID:G79107;
A/Genetics:
A/Genes: KIS

Query Match          47.7%; Score 1052; DB 2; Length 200;
Best Local Similarity 99.0%; Pred. No. 1.3e-53;
Matches 198; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 122 LLYSSHQGGMMWICARDVLEALAPLHHEGYHADKPRNIIISAEENCFKIDFGIS 181
DB 1 LLYSSHQGGMMWICARDVLEALAPLHHEGYHADKPRNIIISAEDECFKIDFGIS 60
QY 182 PKEGNQDVYIOTDGYRAPEALONCLAQAGLSQDTECTSAVDLWSLGIILMFSGMKL 241
DB 61 PKEGNQDVYIOTDGYRAPEALONCLAQAGLSQDTECTSAVDLWSLGIILMFSGMKL 120
QY 242 KHTVRSQEWKANSAAIIDIHFASKAVVNAATPAHYLRDLIKSMHDDPSRRIPAEALACS 301
DB 121 KHTVRSQEWKANSAAIIDIHFASKAVVNAATPAHYLRDLIKSMHDDPSRRIPAEALACS 180
QY 302 PFFSIPAPHIEDLVMLPTP 321
DB 181 PFFSIPAPHIEDLVMLPTP 200

RESULT 2
T24445
hypothetical protein T04C10.1 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T24445
R/Burton, J.
submitted to the EMBL Data Library, March 1996
A/Reference number: Z19891
A/Accession: T24445
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-948 <WIL>

```



```

QY 55 QFLPCTGAAAGAEYGFKEKRALEQLQGHNIYTLGVFTIHSPVNSRC----- 108
DB 57 KLYRPFQSELFAKRA-----YRELRLKHKR-HENVIGLDVFT-----PDETLDDFTD 104
QY 109 --LLELLDVSVSELLLYSHQSCSMIMTOHCARDVLEALAFHHEGYVADLKRNILM 166
DB 105 FYLVMEFMTGDLQKLM---KHETLSBDRIQFLVYOMKIGKYIHAAGVHIRDKPNL-- 159
QY 167 SAENEC-FKLIDRGLSPKESGNODVKYIQTDGYRAPEALONCLAQAGLOSTECTSAVDL 225
DB 160 AVNEDELKLDGLARQADSEMTGYVTRWRPAPEYL-----NMRYTQTVDI 209
QY 226 WSLGILLLENFSGMKLHTVRSQEMKANSALIDHIFASKAVNA----- 270
DB 210 MSVGCIMAEIMITKIL-----FKGN-----DHLQKIMKVTGTPPEFVQKLSA 256
QY 271 -----ALPAYLRD-----LISMLHDPSSRIIPAMALCSFPFSLPRA 309
DB 257 EAKNMEGLPELRKDFASVLTNASPOAVNLKRMVLDEORVTAALAHPEFE----- 312
QY 310 PHIEDLVMLPTVLRLLNVLVDDYDLGNEEYEDVDEVK---EECKK--YGPVVSLLVPR 364
DB 313 -----SLKDTEDPKAQKIDSFDDVDRTELEMKVYKXVLSFKRPPR 355
QY 365 E 365
DB 356 Q 356

```

RESULT 8

```

T45138
protein kinase skp1 [imported] - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C:Accession: T45138
R:Polye, S.E.; Feoktistova, A.; Burke, J.D.; Woodgett, J.R.; Gould, K.L.
Mol. Cell. Biol. 16, 179-191, 1996
A:Title: Schizosaccharomyces pombe skp1+ encodes a protein kinase related to mammalian g
A:Reference number: 229227, MID:96104567, PMID:8524294
A:Accession: T45138
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-354 <P>
A:Cross-references: EMBL:L29449; NID:G1914882; PIDN:AAB51081.1; PID:G1914883
C:Genetics:
A:Gene: skp1
A:Introns: 6/3
C:Function:
A:Pathway: cells w
A>Note: cells with deletions in skp1 are sensitive to heat shock and exhibit defects in
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keyword: phosphoprotein; phosphotransferase
F:335/Binding site: phosphate (Ser) (covalent) #status experimental

```

```

Query Match 10.6%; Score 234.5; DB 2; Length 354;
Best Local Similarity 24.6%; Pred. No. 1.4e-06;
Matches 76; Conservative 47; Mismatches 113; Indels 79; Gaps 10;
QY 29 LGGSSASVYRRCNPGSPGALQFLPRTTGAAGAAAEYGFKEKRAALBQLOGHRN 88
DB 38 VGGSGFVWQVHLIESDK-AAIKRVL-----QDKRFKRELQIKRMIDPN 84
QY 89 IYVLVGVF-----TIHSPNV---PSRCLLELLDVSVSELLLYSHQGS 131
DB 85 IYDLIAYITTGNSDEVYINLVLEFMTYRASRIYTKQKSMPLVYKLV----- 137
QY 132 MMNIOHCARDVLEALAFHHEGYVHADLKERNILMSAENECFKLIDFGLS--FKESNODV 189
DB 138 -----IYQLRSLAVIHAISGICHRDIKPNLLDPENGILKLCDFGSAKILVAGBPV 190
QY 190 KYIQTDGYRAPEALONCLAQAGLOSTECTSAVDLMSLGIILLMEFSG----- 238

```

```

DB 191 SYICSRYYRAPEL-----IFGANDYTHAIDIWSTGCMALMLGHPFPESGID 240
QY 239 -----MKLGHYRSQEMKANSALIDHIFAS-----KAVNAALPAYLRDLKSMLEHD 287
DB 241 QLVETIKIGTSPRSQDKTMNPNVMEHRRPQRLPQPLSVFSRSVSLDL--DLLSKMLQY 299
QY 288 DPSRRIPAMALCSFPF 304
DB 300 TPTDRLTAAMAMCHPFF 316

```

RESULT 9

```

S37790
probable serine/threonine-specific protein kinase (EC 2.7.1.-) YKL161c - yeast (Saccharo
N:Alternate names: protein YKL615
C:Species: Saccharomyces cerevisiae
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 16-Aug-2004
C:Accession: S37790; S37991; S44567
R:Vandenbol, M.; Bolle, P.; Dion, C.; Portetelle, D.; Hilger, F.
submitted to the EMBL Data Library, September 1993
A:Description: DNA sequencing of a 36.2 kb fragment located between the PAS1 and LAP4 loc
A:Reference number: S37786
A:Accession: S37790
A:Molecule type: DNA
A:Residues: 1-433 <VAN>
A:Cross-references: UNIPROT:P36005; EMBL:226877; NID:G407482; PIDN:CAA81493.1; PID:G40748
A:Experimental source: strain S288C
R:Vandenbol, M.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hilger, F.
submitted to the Protein Sequence Database, March 1994
A:Reference number: S37976
A:Accession: S37991
A:Molecule type: DNA
A:Residues: 1-433 <VA2>
A:Cross-references: EMBL:228161; NID:G486280; PIDN:CAA82003.1; PID:G486281; MIPS:YKL161c
A:Experimental source: strain S288C
R:Vandenbol, M.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hilger, F.
Yeast 10, 35-40, 1994
A:Title: DNA sequencing of a 36.2 kb fragment located between the PAS1 and LAP4 loci of c
A:Reference number: S44563
A:Accession: S44567
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-433 <VA3>
A:Cross-references: EMBL:226877; NID:G407482; PIDN:CAA81493.1; PID:G407487
A:Experimental source: strain S288C
C:Genetics:
A:Cross-references: SGD:S0001644
A:Map position: 11L
A:Note: YKL161c
C:Superfamily: protein kinase homology
C:Keyword: ATP; phosphotransferase; serine/threonine-specific protein kinase
F:21-287/Domain: protein kinase homology <KIN>
F:153/Active site: Asp #status predicted

```

```

Query Match 10.6%; Score 234.5; DB 2; Length 433;
Best Local Similarity 29.5%; Pred. No. 1.7e-06;
Matches 89; Conservative 45; Mismatches 97; Indels 71; Gaps 16;
QY 75 KERALELOQGHNIYTLGVFTIHSPVNSRCLL--ELLDVSVSELLLYSHQSGSM 132
DB 71 RELKLRHRLRGHPNIVMLDT--DIVFYNGALNGYVLEHEMCDLSQIT--RSEQRLED 127
QY 133 MMNIOHCARDVLEALAFHHEGYVHADLKERNILMSAENEC-FKLIDFGLS-----FKEG 185
DB 128 AHQPSRIVQILQAKRTIHSAVNLHODIKRNLL--VNSDQGLKICNFGLSGSENNKVN 185
QY 186 NODVK-YIQTDGYRAPEALONCLAQAGLOSTECTSAVDLMSLGIILLF-----MBSG 238
DB 186 DGRIGYIYSIYKAPK-----LINTQECTKAVIDISTGCIILAEELGRRPMFEG 235
QY 239 M-----KLKHTV-----SQEMKANSALIDHIF-----ASKAVNAALPAYHL 277
DB 236 KQYVDHLNHLIQLGTPPETTQD--IASQKYNYIIFQGNIPGSRFSFELIGANP--EA 291

```

QY 278 RDLIKSMHDDPSRRIPAEWALCSPPESI-----PFA-----PHIEDLVMLPTVY 322
 DB 292 LELIKMKLEPDKKRIITVEDALSHPIYSWMHMDIDEEFSCCKTRFEFEHIESMAELGNEY 351
 QY 323 LR 324
 DB 352 IK 353

RESULT 10

S64950
 protein kinase HOG1 (EC 2.7.1.-), mitogen-activated - yeast (Saccharomyces cerevisiae)
 N/Alternate names: protein L2931; protein YNR113W
 C/Species: Saccharomyces cerevisiae
 C/Date: 01-Aug-1995 #sequence_revision 24-May-1996 #text_change 16-Aug-2004
 C/Accession: S64950; S42145; S69400
 R/Vername: P.; Voet, M.; Voelckert, G.
 submitted to the Protein Sequence Database, May 1996
 A/Reference number: S64943
 A/Accession: S64950
 A/Molecule type: DNA
 A/Residues: 1-435 <VER>
 A/Cross-references: UNIPROT:P32485; EMBL:Z73285; NID:g1360507; PDB:CAA97680.1; PID:e245
 A/Experimental source: strain S288C
 R/Brewer: J.L.; de Valoir, T.; Dwyer, N.D.; Winer, E.; Gustin, M.C.
 Science 259, 1760-1763, 1993
 A/Title: An osmoregulating signal transduction pathway in yeast.
 A/Reference number: S42145; MUID:93206121; PMID:7681220
 A/Accession: S42145
 A/Status: nucleic acid sequence not shown
 A/Molecule type: DNA
 A/Residues: 1-408 <GORGSCSK>

 A/Cross-references: EMBL:L06279; NID:g295610; PDB:AAA34680.1; PID:g295611
 R/Vername: P.; Voelckert, G.
 submitted to the EMBL Data Library, September 1995
 A/Reference number: S69393
 A/Accession: S69400
 A/Molecule type: DNA
 A/Residues: 1-435 <VER>
 A/Cross-references: EMBL:X89514; NID:g1297019; PDB:CAA61691.1; PID:e198746; PID:g129702
 C/Genetic: S69393
 A/Genes: SGD:HOG1, SSK3
 A/Cross-references: SGD:S0004103; MIPS:YNR113W
 A/Map position: 12R
 C/Superfamily: protein kinase homology
 C/Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase; signal tr
 F;21-271/Domain: protein kinase homology <KIN>
 F;29-37/Region: protein kinase ATP-binding motif
 F;144/Active site: Asp #status predicted

Query Match 10.6%; Score 234; DB 2; Length 435;
 Best Local Similarity 25.7%; Pred. No. 1.8e-06;
 Matches 98; Conservative 53; Mismatches 140; Indels 90; Gaps 18;

QY 17 EAFGRIMQVQSR-----LGGSSASVYRRCGNNP--GSPGALQFLPPTGTAAGA 68
 DB 11 QIFGVPELTNRNDINPVGMGAFLV---CSATDTLTQPAIKIKMPSTAVLAKR 66
 QY 69 AEGYFKERPALEEQLOGRNIVTLVGVFTIHSPNVPSCILLEDVSVSELLYSHQ 128
 DB 67 T-----YRELKTLKHLR-HENLIGLDQIF---LSP-LIEDIYFVELQGTDLHRL---QTR 114
 QY 129 GCSMMITQICARVLEALAFLLHHEGVHADLKPRNIIWASNECFKILDPGLSPKESND 188
 DB 115 PLEKQVOYQFLYQILGLKTVHSAGVHRDLKPSNIIIN-ENCDLXICDFGLARIODPW 173
 QY 189 VKYIOTDGYAPAEIQLNCLAQAGLSDPTCTSAVDLMSIGITLLMFSGMKL---KHT 244
 DB 174 TGVVSTRYTRAPETML-----TWQKIDV---VDIWSAGCI-PAEMIBSKPLPFGKDH- 222
 QY 245 VRSQEWKANSATIDHIFAS--RAVVAALPAVHLR----- 278

DB 223 -----VHQSIIITDLGSPKQVINTICSENTLKFTVSLPHRDPIDPSEKFKTVEBDA 275
 QY 279 -DLIKSMHDDPSRRIPAEWALCSPPESI-----PFA-----PHIEDLVMLPTVY 322
 DB 276 VDLIKMKLEPDKKRIITVEDALSHPIYSWMHMDIDEEFSCCKTRFEFEHIESMAELGNEY 351
 QY 326 L---NVLDDDYLGNEEVEDV 343
 DB 335 MMYSEILDPHKIKGSDGQIDI 355

RESULT 11

A44878
 protein kinase (EC 2.7.1.37) cdk2 [validated] - goldfish
 C/Species: Carassius auratus (goldfish)
 C/Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
 C/Accession: A44878
 R/Hirai, T.; Yamashita, M.; Yoshikuni, M.; Tokumoto, T.; Kajitara, H.; Sakai, N.; Nagahama
 Dev. Biol. 152, 113-120, 1992
 A/Title: Isolation and characterization of goldfish cdk2, a cognate variant of the cell c
 A/Reference number: A44878; MUID:92331802; PMID:1339336
 A/Accession: A44878
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-298 <HIR>
 A/Cross-references: UNIPROT:P43450; GB:S40289; NID:g251619; PDB:AA822550.1; PID:g251620
 A/Experimental source: oocyte
 A/Note: sequence extracted from NCBI backbone (NCBIN:108782, NCBI:P:108783)
 C/Genetic: S
 A/Genes: cdk2
 C/Superfamily: kinase-related transforming protein; protein kinase homology
 C/Keywords: ATP; cell cycle control; mitosis; phosphoprotein; phosphotransferase; serine/
 F;2-255/Domain: protein kinase homology <KIN>
 F;10-18/Region: protein kinase ATP-binding motif
 F;33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 10.6%; Score 233; DB 1; Length 298;
 Best Local Similarity 28.7%; Pred. No. 1.4e-06;
 Matches 99; Conservative 48; Mismatches 116; Indels 82; Gaps 17;

QY 16 LEAFGRIMQVQSR-----LGGSSASVYRRCGNNPSPGALQFLPPTGTAAGA 68
 DB 1 MSF-----QKVKIKIGEGTVGVYKAK--NKVGTETVALKRIIDTETGVPSTAL----- 49
 QY 75 KERPALEEQLOGRNIVTLVGVFTIHSPNVPSCILLEDVSVSELLYSHQGSCMM 134
 DB 50 RSLISLKEEL-NRPNTVVKHDV--IHTENK---YLVFEFLHODLRRFMDSSVTGTISLPL 103
 QY 135 IOHCARDVLEALAFLLHHEGVHADLKPRNIIWASNECFKILDPGLSPKESNDVYK--- 191
 DB 104 VKSYLEFOLLQGLAFCHSHRVLRDLKPOLNINAGGE-IKLADFLGARAFGVPVATYTHE 162
 QY 192 IOTDGYAPAEIQLNCLAQAGLSDPTCTSAVDLMSIGITLLMFSGMKLKHIVRSQEWK 251
 DB 163 VVTLMYTRAPETML-----KYYSIAVDINSLGCI-PAEMITL-----RKALFP 204
 QY 252 ANSALIDHIFPAKAVN-----AATPAHL----- 280
 DB 205 GSDE--IDQLFRIFRTLTGPDSEIWPGVYSMDYRSPFKARODLSKVPLLEDGRDL 262
 QY 281 IKSMHDDPSRRIPAEWALCSPPESI-----PFA-----PHIEDLVMLPTVY 322
 DB 263 LQGMILYDPNKISAKNALVHRF-----RDVTM-VPVPLRL 298

RESULT 12

A41227
 protein kinase (EC 2.7.1.37) cdk2 - human
 N/Alternate names: Egl homolog; protein kinase p34
 C/Species: Homo sapiens (man)
 C/Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 09-Jul-2004
 C/Accession: A41227; S17873; S16520
 R/Ninomiya-Tsuji, J.; Nomoto, S.; Yasuda, H.; Reed, S.I.; Matsumoto, K.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2004, 11:26:47 ; Search time 122.5 Seconds
(without alignments)
1968.016 Million cell updates/sec

Title: US-10-798-532-2
Perfect score: 2206
Sequence: 1 MAGSGCAGCABPPRFLEAFG.....VATFYPLSAVKRGYLYQTLL 419

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues
Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2198	99.6	419	1	KIST_HUMAN
2	2186	99.1	419	1	KIST_RAT
3	2185	99.0	419	2	AAH58732
4	2185	99.0	419	2	AAO13515
5	2180	98.8	419	1	KIST_MOUSE
6	1711	77.6	330	2	Q8C0N6
7	257	11.7	505	2	Q6CFX3
8	248.5	11.3	882	2	Q8N0L7
9	243.5	11.0	395	2	Q6GM90
10	243.5	11.0	720	2	Q9CA22
11	243.5	11.0	765	2	Q25758
12	243.5	11.0	777	2	Q84VX4
13	243.5	11.0	826	2	Q94656
14	243.5	11.0	914	2	Q811F0
15	242.5	11.0	395	2	Q84YR3
16	242.5	11.0	395	2	AAH70745
17	242	11.0	433	2	Q84SN3
18	241.5	10.9	387	1	SKP1_SCHPO
19	240.5	10.9	360	2	Q6S267
20	240.5	10.9	360	2	AAH17088
21	240	10.9	388	1	SMK1_YEAST
22	238	10.8	298	1	CDK2_MESAU
23	235.5	10.7	434	2	Q75B85
24	235.5	10.7	367	1	MAK1_RAT
25	235.5	10.7	434	2	AA551605
26	234.5	10.6	357	1	MAK1_MOUSE
27	234.5	10.6	352	2	Q62D93
28	234.5	10.6	332	2	BAC99508
29	234.5	10.6	433	1	KKQ1_YEAST
30	234	10.6	298	1	CDK2_CRIGR
31	234	10.6	298	1	CDK2_RAT

32	234	10.6	298	2	Q6P751	Q6P751 rattus norv
33	234	10.6	298	2	AAH61832	AAH61832 rattus no
34	234	10.6	361	1	JNK SUBDO	Q966Y3 subdites d
35	234	10.6	435	1	HOG1_YEAST	P32485 saccharomyc
36	233.5	10.6	433	2	Q6B629	Q6B629 parametium
37	233.5	10.6	1233	2	Q6CAC2	Q6CAC2 yarrowia li
38	233	10.6	298	1	CDK2_CARAU	P43450 carassius a
39	233	10.6	298	1	CDK2_HUMAN	P24941 homo sapien
40	233	10.6	298	2	AAH35467	AAH35467 homo sapi
41	233	10.6	369	2	Q7M445	Q7M445 dictyosteli
42	233	10.6	392	2	Q9YMN3	Q9YMN3 drosophila
43	232.5	10.5	357	2	Q9HD31	Q9HD31 homo sapien
44	232.5	10.5	385	2	Q81U85	Q81U85 homo sapien
45	232.5	10.5	385	2	Q8BW96	Q8BW96 mus musculu

ALIGNMENTS

RESULT 1
KIST_HUMAN
ID KIST_HUMAN STANDARD; PRT; 419 AA.
AC Q8TAS1; Q96C22;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Serine/threonine-protein kinase Kist (EC 2.7.1.37) (Kinase interacting with stahtmin).
GN Name=KIST; Synonyms=KIS;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX PubMed=12782393;
RA Bieche I., Manceau V., Curni P.A., Laurendeau I., Lachkar S., Leroy K., Vidard D., Sobel A., Maucuer A.;
RT "Quantitative RT-PCR reveals a ubiquitous but preferentially neural expression of the KIS gene in rat and human."
RL Brain Res. Mol. Brain Res. 114:55-64(2003).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Eye, and Skin;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares W.B., Bonaldo W.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultky S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman W., Madan A., Rodriguez S., Sanchez A., Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E., Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Possesses kinase activity. May be involved in trafficking and/or processing of RNA (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- SUBUNIT: Interacts with stahtmin (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;

```

CC      IsoId=Q8TAS1-1; Sequence=Displayed;
CC      Name=2;
CC      IsoId=Q8TAS1-2; Sequence=VSP_004908, VSP_004909;
CC      Note=No experimental confirmation available;
CC      -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC      -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AJ536197; CAD60192.1; -.
DR      EMBL; BC014917; AAH14917.1; -.
DR      EMBL; BC026046; AAH26046.1; -.
DR      HSSP; P26368; 100P.
DR      GO; GO:0005737; Cytoplasm; ISS.
DR      GO; GO:0005634; C:nucleus; ISS.
DR      GO; GO:0005524; F:ATP binding; ISS.
DR      GO; GO:0005515; F:protein binding; ISS.
DR      GO; GO:0004674; F:protein serine/threonine kinase activity; ISS.
DR      GO; GO:0003723; F:RNA binding; ISS.
DR      GO; GO:0018105; P:pepckl-1-serine phosphorylation; ISS.
DR      InterPro; IPR011009; Kinase like.
DR      InterPro; IPR000719; Prot kinase.
DR      InterPro; IPR000504; RNA Rec mot.
DR      InterPro; IPR008271; Ser_thr_pkin_AS.
DR      Pfam; PF00669; Kinase; 1.
DR      Pfam; PF00076; RRM_1; 1.
DR      ProDom; PD000001; Prot_kinase; 1.
DR      SMART; SM00360; RRM; 1.
DR      ProDom; PD000001; PROTEIN_KINASE_ATP; FALSE_NEG.
DR      PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR      PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
DR      PROSITE; PS00102; RRM; 1.
DR      PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
DR      Alternative splicing; ATP-binding; Nuclear protein; RNA-binding;
KW      Serine/threonine-protein kinase; Transferase.
FT      DOMAIN 23 304 Protein kinase.
FT      DOMAIN 324 406 RNA-binding (RRM).
FT      NP_BIND 29 37 ATP (By similarity).
FT      BINDING 54 54 ATP (By similarity).
FT      ACT_SITE 158 158 Proton acceptor (By similarity).
FT      VARSPLIC 342 344 DVV -> GIC (in isoform 2).
FT      VARSPLIC 345 419 /FTId=VSP_004908.
FT      VARSPLIC 345 419 Missing (in isoform 2).
FT      VARSPLIC 345 419 /FTId=VSP_004909.
SQ      SEQUENCE 419 AA; 46546 MW; 903E982BE12A8CF8 CRC64;
Query Match 99.6%; Score 2198; DB 1; Length 419;
Best Local Similarity 99.8%; Pred. No. 1.3e-159;
Matches 418; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

DB      241 LKHTVRSQEWKANSALIDHIFASKAVVNAALPAYHLRLIKSLHDDPSRRIPAEMLC 300
QY      301 SPFFSIPFAPHIEDVLMPTPYRLNLNVADDYLGNEEYEDVEDVEKECKRGYVSL 360
DB      301 SPFFSIPFAPHIEDVLMPTPYRLNLNVADDYLGNEEYEDVEDVEKECKRGYVSL 360
QY      361 LVPKENPGRGQVFEVYANAGDSKAAQKLLTGMPDGKPVATFPYPLSAVRYQTLL 419
DB      361 LVPKENPGRGQVFEVYANAGDSKAAQKLLTGMPDGKPVATFPYPLSAVRYQTLL 419
RESULT 2
ID      KIST_RAT STANDARD; PRT; 419 AA.
AC      063285;
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      05-JUL-2004 (Rel. 44, Last annotation update)
DE      Serine/threonine-protein kinase Kist (EC 2.7.1.37) (Kinase interacting
DE      with statmin) (PAM COOH-terminal interactor protein 2) (P-CIP2).
GN      Name=Kist; Synonyms=Kfs;
OS      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_TaxId=10116;
RN      [1]
RP      SEQUENCE FROM N.A., FUNCTION, SUBUNIT, SUBCELLULAR LOCATION, TISSUE
RP      SPECIFICITY, AND MUTAGENESIS OF LYS-54.
RC      TISSUE=Striatum;
RX      MEDLINE=97435279; PubMed=9287318;
RA      Maucuer A., Ozon S., Manceau V., Gavet O., Lawler S., Curmi P.,
RA      Sobel A.;
RT      "Kis is a protein kinase with an RNA recognition motif.";
RL      J. Biol. Chem. 272:23151-23156(1997).
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Caldwell B.D., Darlington D.N., Penzies P., Johnson R.C., Bipper B.A.,
RA      Maine R.E.;
RL      Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE OF 28-419 FROM N.A.
RC      TISSUE=Hippocampus;
RX      MEDLINE=97067094; PubMed=8910496;
RA      Alam R., Caldwell B.D., Johnson R.C., Darlington D.N., Maine R.E.,
RA      Bipper B.A.;
RT      "Novel proteins that interact with the COOH-terminal cytosolic routing
RT      determinants of an integral membrane peptide-processing enzyme.";
RL      J. Biol. Chem. 271:28636-28640(1996).
CC      -1- FUNCTION: Possesses kinase activity. May be involved in
CC      trafficking and/or processing of RNA.
CC      -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC      -1- SUBUNIT: Interacts with statmin (By similarity).
CC      -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.
CC      -1- TISSUE SPECIFICITY: In the embryo, preferentially expressed in the
CC      developing nervous system.
CC      -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC      -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X98374; CAA67021.1; -.
DR      EMBL; U70372; AAC53031.2; -.
DR      HSSP; P26368; 100P.
DR      RGD; 2968; Kistf.
DR      GO; GO:0005737; Cytoplasm; IDA.
DR      GO; GO:0005634; C:nucleus; IDA.

```

DR GO; GO:0005524; F:ATP binding; IC.
 DR GO; GO:0005515; F:protein binding; ISS.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IDA.
 DR GO; GO:0003723; F:RNA binding; NAS.
 DR GO; GO:0018105; P:peptidyl-serine phosphorylation; IDA.
 DR InterPro; IPR011009; Kinase-like.
 DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR000504; RNA-rec mol.
 DR InterPro; IPR008271; Ser-thr-tyr-kin_A5.
 DR Pfam; PF00069; Kinase; I.
 DR Pfam; PF00076; RRM_1; I.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; FALSE_NEG.
 DR PROSITE; PS00111; PROTEIN KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
 DR PROSITE; PS00102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
 DR ATP-binding; Nuclear protein; RNA-binding;
 KM Serine/threonine-protein kinase; Transferase.
 FT DOMAIN 23 304
 FT NP_BIND 324 406
 FT BINDING 29 37
 FT ACT_SITE 141 141
 FT ACT_SITE 158 158
 FT MUTAGEN 54 54
 SQ SEQUENCE 419 AA, 46505 MW, D2F9B7BF8F080DF3 CRC64;

Query Match 99.1%; Score 2186; DB 1; Length 419;
 Best Local Similarity 99.0%; Pred. No. 1.1e-158;
 Matches 415; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAGSGCAGWAEPRFLAEFGRMLQVOSRLGSSSASVYRVRCGPGSPGALKOFLPPG 60
 DB 1 MAGSGCAGWAEPRFLAEFGRMLQVOSRLGSSSASVYRVRCGPGSPGALKOFLPPG 60
 QY 61 TTGAASAABYGFRRERAALEOLQGHRIIVTLVGFTHFSFVNSVRSCLLELDVSVSE 120
 DB 61 TTGAASAABYGFRRERAALEOLQGHRIIVTLVGFTHFSFVNSVRSCLLELDVSVSE 120
 QY 121 LLLYSSHOGCSMMWIOHCARDVLEALFLHESGVHADLKPRNIILMSAENECFKLIDFGL 180
 DB 121 LLLYSSHOGCSMMWIOHCARDVLEALFLHESGVHADLKPRNIILMSAENECFKLIDFGL 180
 QY 121 LLLYSSHOGCSMMWIOHCARDVLEALFLHESGVHADLKPRNIILMSAENECFKLIDFGL 180
 DB 121 LLLYSSHOGCSMMWIOHCARDVLEALFLHESGVHADLKPRNIILMSAENECFKLIDFGL 180
 QY 181 SFKEGNQDVYKIQTDGYRAPEALQNCIAQAGLQSDTECTSAVDLMSLGIILLEMFSGMK 240
 DB 181 SFKEGNQDVYKIQTDGYRAPEALQNCIAQAGLQSDTECTSAVDLMSLGIILLEMFSGMK 240
 QY 241 LKHTVRSQEKANSSAIIDHIFASKAVVNAIPAYHLRDLIKSMHDDPERRIPAEMLC 300
 DB 241 LKHTVRSQEKANSSAIIDHIFASKAVVNAIPAYHLRDLIKSMHDDPERRIPAEMLC 300
 QY 241 LKHTVRSQEKANSSAIIDHIFASKAVVNAIPAYHLRDLIKSMHDDPERRIPAEMLC 300
 DB 241 LKHTVRSQEKANSSAIIDHIFASKAVVNAIPAYHLRDLIKSMHDDPERRIPAEMLC 300
 QY 301 SPFFSIPFAPIEDLVMLPTPVRLINLVDDYLGNEEYEDVDEYVKEECQKYGPVVSL 360
 DB 301 SPFFSIPFAPIEDLVMLPTPVRLINLVDDYLGNEEYEDVDEYVKEECQKYGPVVSL 360
 QY 361 LVPKENPGRQGVFEVYANAGDSKAAQKLLTGEMFGKRVVATFPLSAVYKRGYLQYTL 419
 DB 361 LVPKENPGRQGVFEVYANAGDSKAAQKLLTGEMFGKRVVATFPLSAVYKRGYLQYTL 419

RESULT 3
 AAHS8732
 ID AAHS8732 PRELIMINARY; PRT; 419 AA.
 AC AAHS8732;
 DT 02-MAR-2004 (TrEMBLrel. 27, Created)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
 DE Kinase interacting with leukemia-associated gene (Statmin).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strusberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Burow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diachenko L., Marutina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleorn M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Frange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Faney J., Helton B., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butcherfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.,
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RA Strusberg R.,
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC058732; AAHS8732.1; -.
 KW Kinase.
 SQ SEQUENCE 419 AA, 46489 MW, 3BD3C06A59E22540 CRC64;

Query Match 99.0%; Score 2185; DB 2; Length 419;
 Best Local Similarity 99.0%; Pred. No. 1.3e-158;
 Matches 415; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAGSGCAGWAEPRFLAEFGRMLQVOSRLGSSSASVYRVRCGPGSPGALKOFLPPG 60
 DB 1 MAGSGCAGWAEPRFLAEFGRMLQVOSRLGSSSASVYRVRCGPGSPGALKOFLPPG 60
 QY 61 TTGAASAABYGFRRERAALEOLQGHRIIVTLVGFTHFSFVNSVRSCLLELDVSVSE 120
 DB 61 TTGAASAABYGFRRERAALEOLQGHRIIVTLVGFTHFSFVNSVRSCLLELDVSVSE 120
 QY 61 TTGAASAABYGFRRERAALEOLQGHRIIVTLVGFTHFSFVNSVRSCLLELDVSVSE 120
 DB 61 TTGAASAABYGFRRERAALEOLQGHRIIVTLVGFTHFSFVNSVRSCLLELDVSVSE 120
 QY 121 LLLYSSHOGCSMMWIOHCARDVLEALFLHESGVHADLKPRNIILMSAENECFKLIDFGL 180
 DB 121 LLLYSSHOGCSMMWIOHCARDVLEALFLHESGVHADLKPRNIILMSAENECFKLIDFGL 180
 QY 121 LLLYSSHOGCSMMWIOHCARDVLEALFLHESGVHADLKPRNIILMSAENECFKLIDFGL 180
 DB 121 LLLYSSHOGCSMMWIOHCARDVLEALFLHESGVHADLKPRNIILMSAENECFKLIDFGL 180
 QY 181 SFKEGNQDVYKIQTDGYRAPEALQNCIAQAGLQSDTECTSAVDLMSLGIILLEMFSGMK 240
 DB 181 SFKEGNQDVYKIQTDGYRAPEALQNCIAQAGLQSDTECTSAVDLMSLGIILLEMFSGMK 240
 QY 241 LKHTVRSQEKANSSAIIDHIFASKAVVNAIPAYHLRDLIKSMHDDPERRIPAEMLC 300
 DB 241 LKHTVRSQEKANSSAIIDHIFASKAVVNAIPAYHLRDLIKSMHDDPERRIPAEMLC 300
 QY 241 LKHTVRSQEKANSSAIIDHIFASKAVVNAIPAYHLRDLIKSMHDDPERRIPAEMLC 300
 DB 241 LKHTVRSQEKANSSAIIDHIFASKAVVNAIPAYHLRDLIKSMHDDPERRIPAEMLC 300
 QY 301 SPFFSIPFAPIEDLVMLPTPVRLINLVDDYLGNEEYEDVDEYVKEECQKYGPVVSL 360
 DB 301 SPFFSIPFAPIEDLVMLPTPVRLINLVDDYLGNEEYEDVDEYVKEECQKYGPVVSL 360
 QY 361 LVPKENPGRQGVFEVYANAGDSKAAQKLLTGEMFGKRVVATFPLSAVYKRGYLQYTL 419
 DB 361 LVPKENPGRQGVFEVYANAGDSKAAQKLLTGEMFGKRVVATFPLSAVYKRGYLQYTL 419

RESULT 4
 AA013515
 ID AA013515 PRELIMINARY; PRT; 419 AA.
 AC AA013515;
 DT 02-MAR-2004 (TrEMBLrel. 27, Created)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)

DT 02-MAR-2004 (TEMBLrel. 27, Last annotation update)
 DE KIS kinase.
 GN KIS.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=12956;
 RA Crook M.F., Boehm M., Nabel E.G.;
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY180177; AA013515.1; -.
 SK Kinase.
 SQ SEQUENCE 419 AA; 46489 MW; 3BD3C06A59E22540 CRC64;
 Query Match 99.0%; Score 2185; DB 2; Length 419;
 Best Local Similarity 99.0%; Pred. No. 1.3e-158; Indels 0; Gaps 0;
 Matches 415; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 MAGSGCAGMAGPPRFLFAFGRILWQVOSRLGSGSASVYRVRCGNGPPGALKQFLPPG 60
 1 MAGSGCAGMAGPPRFLFAFGRILWQVOSRLGSGSASVYRVRCGNGPPGALKQFLPPG 60
 Db
 QY 61 TTGAASAAAYGRKRAALAEQAGHNITLYGVFTIHSPVPSKCLLELDVSVSE 120
 61 TTGAASAAAYGRKRAALAEQAGHNITLYGVFTIHSPVPSKCLLELDVSVSE 120
 Db
 QY 121 LLLYSSHQGSMWIOHCARDVLEALAFLEHGVHADLPRIILMSAENECFGLIDFGL 180
 121 LLLYSSHQGSMWIOHCARDVLEALAFLEHGVHADLPRIILMSAENECFGLIDFGL 180
 Db
 QY 121 LLLYSSHQGSMWIOHCARDVLEALAFLEHGVHADLPRIILMSAENECFGLIDFGL 180
 121 LLLYSSHQGSMWIOHCARDVLEALAFLEHGVHADLPRIILMSAENECFGLIDFGL 180
 Db
 QY 181 SFPEGNQDVYKIQTDGYRAPEAEIQLNCLAQGLQSDTECTSAVDLNSLGIILLEMESGMK 240
 181 SFPEGNQDVYKIQTDGYRAPEAEIQLNCLAQGLQSDTECTSAVDLNSLGIILLEMESGMK 240
 Db
 QY 181 SFPEGNQDVYKIQTDGYRAPEAEIQLNCLAQGLQSDTECTSAVDLNSLGIILLEMESGMK 240
 181 SFPEGNQDVYKIQTDGYRAPEAEIQLNCLAQGLQSDTECTSAVDLNSLGIILLEMESGMK 240
 Db
 QY 241 LKRTVNSQEKANSALIIDHIFASKAVVNAIPAYHRLDKMLHDDPSRRIPAEALC 300
 241 LKRTVNSQEKANSALIIDHIFASKAVVNAIPAYHRLDKMLHDDPSRRIPAEALC 300
 Db
 QY 241 LKRTVNSQEKANSALIIDHIFASKAVVNAIPAYHRLDKMLHDDPSRRIPAEALC 300
 241 LKRTVNSQEKANSALIIDHIFASKAVVNAIPAYHRLDKMLHDDPSRRIPAEALC 300
 Db
 QY 301 SPFFSIPFAPIHEDLVMLPTPVRLNLNVDDDLGNNEEVEDVEDVEKCECKYGPVSL 360
 301 SPFFSIPFAPIHEDLVMLPTPVRLNLNVDDDLGNNEEVEDVEDVEKCECKYGPVSL 360
 Db
 QY 301 SPFFSIPFAPIHEDLVMLPTPVRLNLNVDDDLGNNEEVEDVEDVEKCECKYGPVSL 360
 301 SPFFSIPFAPIHEDLVMLPTPVRLNLNVDDDLGNNEEVEDVEDVEKCECKYGPVSL 360
 Db
 QY 361 LVPEKENGRCQVFEVYANAGDSKAAQGLTGRMFQKFFVATFPLSAVYRGYLYQTL 419
 361 LVPEKENGRCQVFEVYANAGDSKAAQGLTGRMFQKFFVATFPLSAVYRGYLYQTL 419
 Db
 QY 361 LVPEKENGRCQVFEVYANAGDSKAAQGLTGRMFQKFFVATFPLSAVYRGYLYQTL 419
 361 LVPEKENGRCQVFEVYANAGDSKAAQGLTGRMFQKFFVATFPLSAVYRGYLYQTL 419
 Db

RESULT 5

KIST MOUSE

ID KIST MOUSE STANDARD; PRT; 419 AA.
 AC P97343; G61775; O9CYTL;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Serine/threonine-protein kinase Kist (EC 2.7.1.37) (Kinase interacting
 with statmin).
 GN Name=Kist; Synonyms=KIS;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaido I., Oseba N., Saito R., Suzuki H., Yamataka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T.,
 RA Balderelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schirml L.M., Kampin A., Matsuda H., Batalov S., Batzel K.W.,
 RA Blake J.A., Bradt D., Bruste V., Chochia C., Cordan L.B., Cousins S.,

RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimm S., Guenrich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Kanagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglocz D.R., Malais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.D., Perera G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sanderlin A., Schneider C., Sempé C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wyszewski B., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arikawa T., Fukuda S.,
 RA Hara A., Hashitume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RA "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs.";
 RL Nature 420:563-573 (2002).
 RN (2)
 RP SEQUENCE OF 1-414 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=97435279; PubMed=9287318;
 RA Maucuer A., Ozon S., Manceau V., Gavet O., Lawler S., Curmi P.,
 RA Sobel A.;
 RT "Kis is a protein kinase with an RNA recognition motif.";
 RL J. Biol. Chem. 272:23151-23156(1997).
 RN (3)
 RP REVISIONS TO 52 AND 68-69.
 RA Maucuer A.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 RN (4)
 RP SEQUENCE OF 122-321 FROM N.A., AND INTERACTION WITH STATMIN.
 RC TISSUE=Embryo;
 RX MEDLINE=95241452; PubMed=7724523;
 RA Maucuer A., Camonis J.H., Sobel A.;
 RT "Statmin interaction with a putative kinase and coiled-coil-forming
 protein domains.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:3100-3104(1995).
 CC -1- FUNCTION: Possesses kinase activity. May be involved in
 CC trafficking and/or processing of RNA (By similarity).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphorylated protein.
 CC -1- SUBUNIT: Interacts with statmin.
 CC -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Y10725; CAA71714.2; -.
 DR EMBL: AK013347; BAB28802.1; -.
 DR EMBL: X62320; CAA57763.1; -.
 DR PIR: I48615; I48615.
 DR HSSP: P26368; 100P.
 DR MGD: MGI:1341908; Kist.
 DR GO: GO:0005737; Cytoplasm; ISS.
 DR GO: GO:0005634; Cytoplasm; ISS.
 DR GO: GO:0005524; F-ATP binding; ISS.
 DR GO: GO:0005515; F-protein binding; IPT.
 DR GO: GO:0004674; F-protein serine/threonine kinase activity; ISS.
 DR GO: GO:0003723; F-RNA binding; ISS.
 DR GO: GO:0018105; Peptidyl-serine phosphorylation; ISS.
 DR InterPro: IPR011009; Kinase_like.

DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR Pfam; PF00069; Pkinase; 1.
 DR Pfam; PF00076; RRM_1; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; FALSE_NEG.
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
 DR ATP-binding; Nuclear protein; RNA-binding;
 KM Serine/threonine-protein kinase; Transferase.
 FT DOMAIN 23 303
 FT BINDING 29 37 RNA-binding (RRM).
 FT BINDING 54 54 ATP (By similarity).
 FT ACT SITE 158 158 ATP (By similarity).
 FT CONFLICT 170 170 Proton acceptor (By similarity).
 FT CONFLICT 231 231 N -> D (in Ref. 4).
 SQ SEQUENCE 419 AA; 46477 MW; A91B307BA50F29A1 CRC64;

Query Match 98.8%; Score 2180; DB 1; Length 419;
 Best Local Similarity 98.8%; Pred. No. 3.2e-158;
 Matches 414; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 MASSGCMGAEPPRFLEAFGRILNQVSRGLSSGSASTYRRRCGCGPPGALKQFLPPG 60
 1 MASSGCMGAEPPRFLEAFGRILNQVSRGLSSGSASTYRRRCGCGPPGALKQFLPPG 60

DB 1 TTGAASAAEYGRKERAALEQGHNIYTLGVFTIHSPVPSRCLLELLDVSE 120
 61 TTGAASAAEYGRKERAALEQGHNIYTLGVFTIHSPVPSRCLLELLDVSE 120

QY 121 LLLYSSHOGCSMMWIOHCAVDLEALFLHHEGVNADLPRNITLSAENECEFLIDPGL 180
 121 LLLYSSHOGCSMMWIOHCAVDLEALFLHHEGVNADLPRNITLSAENECEFLIDPGL 180

DB 121 LLLYSSHOGCSMMWIOHCAVDLEALFLHHEGVNADLPRNITLSAENECEFLIDPGL 180
 121 LLLYSSHOGCSMMWIOHCAVDLEALFLHHEGVNADLPRNITLSAENECEFLIDPGL 180

QY 181 SFPEGNDVXYITQDGYRAPEALQNLQAGLOSTRECTSAVDWLSGLITLLEMSGMK 240
 181 SFPEGNDVXYITQDGYRAPEALQNLQAGLOSTRECTSAVDWLSGLITLLEMSGMK 240

DB 181 SFPEGNDVXYITQDGYRAPEALQNLQAGLOSTRECTSAVDWLSGLITLLEMSGMK 240
 181 SFPEGNDVXYITQDGYRAPEALQNLQAGLOSTRECTSAVDWLSGLITLLEMSGMK 240

QY 241 LKHTVRSQEWKASSATIDHIFASKAVVNAIPYHLRDLKSLHDDPSRRIPAEALC 300
 241 LKHTVRSQEWKASSATIDHIFASKAVVNAIPYHLRDLKSLHDDPSRRIPAEALC 300

DB 241 LKHTVRSQEWKASSATIDHIFASKAVVNAIPYHLRDLKSLHDDPSRRIPAEALC 300
 241 LKHTVRSQEWKASSATIDHIFASKAVVNAIPYHLRDLKSLHDDPSRRIPAEALC 300

QY 301 SPFFSIPFAHIEDLVLPVRLINLVDDYLGNEEYEDVEDYKEECOKYGPVSL 360
 301 SPFFSIPFAHIEDLVLPVRLINLVDDYLGNEEYEDVEDYKEECOKYGPVSL 360

DB 301 SPFFSIPFAHIEDLVLPVRLINLVDDYLGNEEYEDVEDYKEECOKYGPVSL 360
 301 SPFFSIPFAHIEDLVLPVRLINLVDDYLGNEEYEDVEDYKEECOKYGPVSL 360

QY 361 LVPEKENGQGVVEYVNAAGDSKAAQKLTGRMGDKFVATVPYPSAYRGYLYQTLL 419
 361 LVPEKENGQGVVEYVNAAGDSKAAQKLTGRMGDKFVATVPYPSAYRGYLYQTLL 419

DB 361 LVPEKENGQGVVEYVNAAGDSKAAQKLTGRMGDKFVATVPYPSAYRGYLYQTLL 419
 361 LVPEKENGQGVVEYVNAAGDSKAAQKLTGRMGDKFVATVPYPSAYRGYLYQTLL 419

RESULT 6
 Q8CON6 PRELIMINARY; PRT; 330 AA.

AC Q8CON6;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Mus musculus adult male testis cDNA, RIKEN full-length enriched
 DE library, clone:493401D07 product:kinase interacting with leukemia-
 DE associated gene (statmin), full insert sequence.
 GN Name=Kist;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=92279253; PubMed=10349636;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=21085660; PubMed=11217851;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RA The FANTOM Consortium;
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=2049374; PubMed=11042159;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=20530913; PubMed=11076861;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Saeki N., Carninci P.,
 RA Kono H., Akiyama J., Nishi K., Katsunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwara S., Inoue K., Ozawa Y., Izawa M., Ohara E., Wachihi M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Marsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RA Adachi J., Aizawa K., Akimura T., Arikawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoaka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katori H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi P., Takaku-Akahira S., Tanaka Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK030152; BAC26812.1; -.
 DR HSSP; P26368; 100P.
 DR MGD; MGI:1341908; Kist.
 DR GO; GO:0005224; F-ATP binding; IEA.
 DR GO; GO:0004672; F-protein kinase activity; IEA.
 DR GO; GO:0016740; F-transferase activity; IEA.
 DR GO; GO:0006468; P-protein amino acid phosphorylation; IEA.
 DR InterPro; IPR011009; Kinase like.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00069; Pkinase; 2.
 DR Pfam; PF00076; RRM_1; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

KR		PROSITE, PS50102; RRM; 1.
RN		KLINBASE
SO	SEQUENCE	330 AA; 37201 MW; 4164C0758BE7B4C CRC64;
OY		Query Match 77.6%; Score 1711; DB 2; Length 330; Best Local Similarity 98.8%; Pred. No. 1..7e-122;
Dc	MATCHES	326; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
OY	90	VTLGVFTITHSNPVPSKCLLELDVDVSSELLYSHQGSGMMIQCARDYLEALFL 149 :
Dc	1	MTLVGFPIHTSPNSPVRSCLLIELLIDVASSELILYSSHGQSGMNIQHCAADVLENALFL 60
OY	150	HHEGGVAHDLPRNTLMASANECPGLDPFGUSFKXGNDDVKYIQTDGYRAPAELOCLA 209
Dc	61	HHEGVPAHDLPKRNLTMASENECFKLIPDGUSFKXGNDDVKYIQDGRAPAELOCLA 120
OY	210	QAQLSDPTRECTSAVDLMSLGILLFEMPGMKLKHTVRQEOWKANSSAIITHIFASKAVVN 269
Dc	121	QAQLSDPTRECTSAVDLMSLGILLFEMPGMKLKHTVRQEOWKANSSAIIIDHFASKAVVN 180
OY	270	AAlPAYHLRLDIKSMLHDDPSRRIPAENALCSPPFSIPPAPHIEDVLVPPLVLRLLNVLT 329
Dc	181	AAIPAYHRDRDKSMTHADDPCRRIIPAEAMALCSPPFSIPPAPHIEDVLWLPVLRLLNVLT 240
OY	330	DDDYLNGEEXEEEDVEDVEDEKEOQQKGPVSVSLYPKENRGROVFPEVANAGDSKAOKTLL 389
Dc	241	DDDYVENDEYEDEDVEDVEDECOQRGPVSVSLYPKENRGROVFPEVANAGDSKAOKTLL 300
OY	390	TGRMPDGKFVVATFPYLAYSARKGYLYQTLL 419
Dc	301	TGRMPDGKFVVATFPYLAYSARKGYLYQTLL 330
RESULT 7		
ID	OCFCFX3	PRELIMINARY; PRT; 505 AA.
AC	OCFCFX3	:
DT	01-OCT-2004 (TREMBLrel. 28, Created)	
DT	01-OCT-2004 (TREMBLrel. 28, Last sequence update)	
DT	01-OCT-2004 (TREMBLrel. 28, Last annotation update)	
DE	Chromosome B of strain CLIB99 of Yarrowia lipolytica.	
CN	ORFNAMES=YALI0B02816g;	
OS	Yarrowia lipolytica (Candidate lipolytica).	
CC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;	
OC	Saccharomycetales; Dipodascaceae; Yarrowia.	
OX	NCB1_TaxID=4952;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=CLIB99;	
RG	GEOLEUVRES;	
RA	Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S., LaFontaine I., de Montigny J., Marcq C., Neuveglise C., Talla E., Goiffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V., Barney S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C., Boisserie A., Boyer D., Catcollito U., Confalonieri F., de Darvarat A., Despons L., Fabre E., Fairhead C., Ferry-Dunazet H., Groppi A., Hantraire F., Hennequin C., Jauniaux N., Joyet P., Kachouri R., Ra Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Mullier H., Nicand J.M., Nikolski M., Oztas S., Ozler-Kalogeropoulous O., Pelencz S., Potier S., Richard G.F., Strub M.L., Suleau A., Swenneze D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B., Zeilon-Meyer M., Zivanovic I., Bolotin-Pokutara M., Thierry A., Bouchier C., Caubron B., Scarpelli C., Galliard C., Weissenbach J., Wincker P., Soucié J.L.; "Genome evolution in yeasts." Nature 430:35-44(2004). [2]	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=CLIB99;	
RA	GENSCOPE:	
RL	Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases. EMBL, CR362128, CAG82657.1, --	
SO	SEQUENCE	505 AA; 58191 MW; D30933013FCCE357 CRC64;

	Query Match	11.7%	Score 257;	DB 2;	Length 505;	
	Best Local Similarity	25.5%	Pred. No. 4.5e-11;			
	Matches	88;	Conservative	46;	Mismatches 122;	Indels 82; Gaps 11
QY	23	WQVSRLLGSGSSASAYRVRCGNGPSPGALKOFLPPGTGTAAMASAEYGRKERALLQ	82			
DB	22	FQVTRKEIGHGAYGVYCAKAYGTGPDGFOVALKKVTNIPSKNIIICKRA---	77			
QY	83	LOGHRNIVTLXGVFTIHFSFNVPSRCLLELLDV-SVELLEYSHQGCSMMI-----	135			
DB	78	FRGHKNITICLDV-----MDIVDTNNFNELELYELMECDNHQIIRSGOP	121			
QY	136	-----OHCADVLEALFLHHEGVNADLKPRNIIWMAENECFLIDPFGSPKEGNQDV	189			
DB	122	LTDAAHYSFTVQIILAGVKIKHSADVLHDLKPGNIIWVANADE-LKICDPLGARGSSDDE	180			
QY	190	K-----YIOTDGYRAPEALQNCIAQAGLSQDTECTSAVDLWSLGIILLEMFSGMKLK	242			
DB	181	KNAGELTEYVATRMRYRAPEIWL-----SPGSYTKAIDIMVSGCIIAMELLGG---K	227			
QY	243	HTVRSQEWKANSASAIIDHI-----FASKAVVNAAIPAYH--	276			
DB	228	PLFGKKNVVDQNOILHLYGTSPSEETLRIGSPRAQBYVRGLPFPMPKIPFSTLPTANPE	287			
QY	277	LRLDIKSLHDDPSRRIPAEAMALCSPFSPSIFPAPIHEDLVNLPMP	321			
DB	288	ALDLERKALDPARVTVVEALTEHPYKTIWHD--RDEVCPTD	330			

ID	ORGANISM	PRELIMINARY	PROT	882 AA
AC	OSMOL7	Q22155;		
DT	01-MAR-2002	(TrEMBLrel. 20, Created)		
DT	01-MAR-2002	(TrEMBLrel. 20, Last sequence update)		
DT	01-OCT-2004	(TrEMBLrel. 28, Last annotation update)		
DE	Mitochondrial kinase (Hypothetical protein T04C10.1).			
GN	Name=mbk-1; Synonyms=T04C10.1;			
OS	Caenorhabditis elegans.			
OC	Eukaryotes; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;			
OC	Rhabditidae; Peldoderinae; Caenorhabditis.			
OX	NCBI_TaxID=6239;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=22505392; PubMed=12618396;			
RA	Rauch W.B., Mootman C., Laceyfield C.O., Leher J., Bartech D.,			
RA	Plasterk R.H., Kandel E.R., Hobert O.;			
RT	"Characterization of Caenorhabditis elegans homologs of the Down			
RT	syndrome candidate gene DYRK1A.";			
RL	Genetics 163:571-580(2003).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Br1sc01 N2;			
RX	MEDLINE=99069613; PubMed=9851916;			
RA	none;			
RT	"Investigating biology.";			
RT	"Investigating biology.";			
RL	Science 282:2012-2018(1998).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Br1sc01 N2;			
RA	Burton J.;			
RL	Submitted (MAR-1996) to the EMBL/Genbank/DBJ databases.			
CC	-I- SIMILARITY: Belongs to the Ser/Thr protein kinase family.			
DR	EMBL; AY064464; AAL40874.1; --			
DR	EMBL; Z69885; CA93756.2; --			
DR	HSSP; P24941; 1H0V.			
DR	Interact; OSMOL7; --			
DR	GO; GO:0005524; F:ATP binding; IEA.			
DR	GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.			
DR	GO; GO:0016740; F:transferase activity; IEA.			
DR	GO; GO:0004668; P:protein amino acid phosphorylation; IEA.			

DR InterPro: IPR011009; Kinase like.
DR InterPro: IPR000719; Prot kinase.
DR InterPro: IPR002290; Ser Thr kinase.
DR InterPro: IPR008271; Ser Thr_kin_AS.
DR Pfam: PF00069; Pkinase; 1.
DR ProDom: PD000001; Prot kinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Hypothetical protein; Kinase;
KW Serine/threonine-protein kinase; Transferase;
SQ SEQUENCE 882 AA; 101244 MW; DDC9DD354DA709 CRC64;

Query Match 11.3%; Score 248.5; DB 2; Length 882;
Best Local Similarity 25.7%; Pred. No. 4e-10;
Matches 104; Conservative 57; Mismatches 113; Indels 131; Gaps 17;
QY 88 NIVTLGVFTIHPSPVNSRCLLLELDVSVSELLYSHQSCMMWIGHCARDVLEALA 147
DB 389 NIVTLKGHP-VHRA-----HCLVPELLSYLVYDLKNTSRGVSLNARKFAQGLKTL 443
QY 148 FLHHE--GVVHADLKRNILW-SAENECEFLIDGLSFKEGNODVXYIQTDGYRADEAL 204
DB 444 FLSSPELSTIHCDELKPEVNLVNAKRSQIRVIDFGSSCGTHRIYQYIGRFYRSEVLL 503
QY 205 QNCLAQAGLODPECHSAVDLWNLGILLLEMEG-----MLQ----- 241
DB 504 -----GLAYDYK-----IDWMLGCLVEMHAGEPLFAGSSVDQMKEVYLGMPKKE 552
QY 242 -----KHTVRSQSWKNSAIIIDHIFASRAVNAAI 272
DB 553 MLDIGKRTKHYPKTDEGIYCKKTDGYRHT-----YARPKAKLHEILG-----VTSGG 603
QY 273 PA-----YHLRDLIKSLMHDPSRIRPAEMALCSPPFSI--PFAPIEDLV 316
DB 604 PGRRLGEPGSHVEDYSKFDLIKRLQFDPKORI-----SPYVVRHPLFKQKEERY 656
QY 317 MLPTPLRLNLVDDDYLGNEEVEEDVDEYKECKY--GPRVSLVLPENGRQGVY 374
DB 657 PGPVPVSH-----SNLQOQOOLXYIQPSQMSQVNEPSVGSVYV 696
QY 375 E-----YANAGDSKAAQGLTGRMFDGKF--VVAIFYPLASIKKGY 413
DB 697 EDNGMTRQARFGSSANPISVTSSFDGDAMEVDAGRRRRFSHQOY 741

RESULT 9

Q6GM90 PRELIMINARY; PRT; 395 AA.
AC 06GM90;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
NCBI_TaxID=8355;
RN 1} SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Hsieh F.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marinsina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Tohbiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loguelli N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shcherchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN 12}
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22241132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.,
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative";
RL Dev. Dyn. 225:384-391 (2002).
RN 13}
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Klein S., Strausberg R.;
RA Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL: BC074183; AH74183.1; -
DR InterPro: IPR011009; Kinase like.
DR InterPro: IPR000719; Prot kinase.
DR InterPro: IPR002290; Ser Thr kinase.
DR InterPro: IPR008271; Ser Thr_kin_AS.
DR InterPro: IPR001245; Tyr kinase.
DR Pfam: PF00069; Pkinase; 1.
DR ProDom: PD000001; Prot kinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR SMART: SM00219; TYRK; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Hypothetical protein; Kinase;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 395 AA; 43896 MW; 4C78BB8F828A2CC4 CRC64;
Query Match 11.0%; Score 243.5; DB 2; Length 395;
Best Local Similarity 27.7%; Pred. No. 3.5e-10;
Matches 91; Conservative 47; Mismatches 108; Indels 83; Gaps 16;
QY 4 SGCWGAEPPEPFLFAFGRLMOVQSRIGSGSSASV-----YRVCCGNPGSPGA 52
DB 9 SGAQWKKQ-----VEDIKRMFQKEVLTGAFSEVYLAEEKETGKFAVACI-----PKKA 59
QY 53 LKQFLPFGTTGAAGAAGYGFKEKPALEFQLOGHRIYTLVGVFTIHPSPVNSRCLL- 111
DB 60 LK-----GKXSEIENEIAVLRIKIR-HENIVALEIDY-----ESPSSLYLW 99
QY 112 -----ELIDVSVSELLYSHQSCMMWIGHCARDVLEALFLHHEGVHADLKRNIL 165
DB 100 QLVSGGEELPD-RIVEKGFTEKQASTL-----HQVLDVAVSYLRILGIVHADLKRNIL 152
QY 166 WSAENECEFLI--DFGLSFKEGNODVXYIQ--TDGYRAPEALQNCIAQAGLSDECT 221
DB 153 YFSQVBEASKIMTSDGLSKMEKGDVMSYACGTPGVAAPEV-----LQ-----KPYSK 201
QY 222 AVDLWNLGILLLEMEBSGMLKKTIVRSQEKANKSSAIIIDHIFASRAVNAAIAY-----H 276
DB 202 AVDCWSIGIVAYITLLCGY-----PPFYDENDSRLFEQIT--KADYEPDIPYWDISBS 252
QY 277 LRDLIKSLMHDPSRIRPAEMALCSPPFS 305
DB 253 ADFQIONLEKQDPNKKYTCQALRHFWIA 281

RESULT 10

09CA22
AC 09CA22 PRELIMINARY; PRT; 720 AA.
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE Hypothetical protein T3288.5.
GN Name=T3288.5;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid1; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lin X., Kaul S., Town C.D., Bentto M., Creasy T.H., Haas B.J., Wu D.,
RA Maizl R., Roming C.M., Koo H., Fujii C.Y., Uteback T.R.,
RA Barnstead M.E., Bowman C.L., White O., Nieman W.C., Fraser C.M.,
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Town C.D., Kaul S.,
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC012193; MAG51619.1; -.
DR PIR; A96807; A96807.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:Protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR000719; Prot_Kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00069; Pkinase; 1.
DR Prodom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE; PS00108; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
DR ATP-binding; Hypothetical protein; Kinase; Transferase.
KW SEQUENCE 720 AA; 80027 MW; F26A9A0B1C6D91DE CRC64;
SQ

Query Match 11.0%; Score 243.5; DB 2; Length 720;
Best Local Similarity 25.9%; Pred. No. 7.5e-10;
Matches 106; Conservative 56; Mismatches 146; Indels 101; Gaps 17;

11 EPPRFLEAFGRLLQVQSRSLGSSASVYRVCCGPPGALKQFLPPTGAAASAAE 70
DB 331 DPDLFRKVNCKLQRLGKISSGSSSEVHKV---SSDCTYALKTKLKGKRDYATA---- 383
QY 71 YGFRKRAALEQIQGHRIIVTLVGVFTIHPSPNVPSRCLLELDVSVSE----- 120
DB 384 YGFCQELIGVLLKKGKTNIIQL-----IDVEVTDKTLLEGLVNGTMSKDGKVEDGF 436
QY 121 ---LLIYS-----SHQCSMW-----MICHGARDVLEALFLHHEGYVADLK 160
DB 437 IYVAVLEYGEIDLAMHSQKWRLEIGSDRTIDENWLEFYMQOIIQAVNTIHEERIVASDK 496
QY 161 PRNIIISAENEFCRLIDFGSPKESQODVXYIOTD-----GYRAEALONCLAQAGQ 214
DB 497 PANFL--LVRGFKLIDFGIA-KAINSDTNIGRDSQVGTLSWSPALFACNESDENG-- 551
QY 215 SDTECTSAVDLWSLGIILLEMFGSKMLKATVRSQEWKANSAAIID--HIFASKAVVNAAI 272
DB 552 NIKCGRPSDWSLGLCYQWVG-RTPFADYKTFW-AKFKVITDNNHETTYQLSNP-- 607
QY 273 PAYHLDLISMLHDDSRRIIPAEMLCSPFESIIPAPH----- 311
DB 608 ---WLDLMKCKCLAWRNQWRRIPELLQHPFLAPPI-PHEPQVKTIKLPSLIAESGSD 663
QY 312 -----IEDLVMLPTVLRLLNVLDLDYLGNEEEYEDVADVEKEC 351
DB 664 DKANSMSIQLEQLSNPAPLPR-----NDVLSDRDQNOQLSRVSELC 706

RESULT 11
ID 025758 PRELIMINARY; PRT; 765 AA.
AC 025758;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE Mitogen-activated protein kinase-related protein.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Dd2;
RC Lin D.T., Goldman N.D., Syin C.;
RT "Stage-specific expression of a Plasmodium falciparum protein related
RL to the eukaryotic mitogen-activated protein kinases";
RN Mol. Biochem. Parasitol. 78:0-0(0).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=Dd2;
RC Lin D.T., Goldman N.D., Syin C.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; U63377; AAC47170.1; -.
DR HSSP; Q16539; 1KVI.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:000468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR003527; MAP_kin.
DR InterPro; IPR000719; Prot_Kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00069; Pkinase; 1.
DR Prodom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS01351; MARK; UNKNOWN_1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
KW SEQUENCE 765 AA; 90231 MW; 5BEC8E6A4B02D32 CRC64;
SQ

Query Match 11.0%; Score 243.5; DB 2; Length 765;
Best Local Similarity 25.8%; Pred. No. 8.1e-10;
Matches 105; Conservative 58; Mismatches 145; Indels 99; Gaps 20;

23 WQVSRSLGSSASVYRVCCGPPGALKQFLPPTGAAASA--AEYGRKRAAL 80
DB 23 YDLKKVKGAGVGFYKGRCKKNKIY--AVYKIF-----GAFQCTDAGRFR-EIIFL 74
QY 81 EQLQGHRIIVTLVGVFTIHPSPNVPSRCLLELDVSVSEL---LLYSHQCSMMMIQ 136
DB 75 YELNGHDNIITKLMDYIK--AKNDNDIYILDPFMTDLDEVIKADLLLEIHK-----K 124
QY 137 HCARVLEALFLHHEGYVHADLKPRTILMSLENKC-FKLIDFGIS-----FKESQNDV 189
DB 125 YIIYQLRLKTYIHSGGLHRDIKPSNLI--VNSCHIKVADFGIARISITHVANKVPI 182
QY 190 --KYIQTDGVARPEALEQNCIAQAGLSPTECTSAVDLWSLGIILLEMFGSKML---KHT 244
DB 183 LRDYVATRYRARPFI-----LGSHTYTEDVDWMSLGCINGELLCGRFLPTGNSI 232
QY 245 VRSQE-----WKANSAAIID--HIFASKAVVN-AAIPAYHLR-----DLIKSM 284
DB 233 KMQLEKIIQIVGKPKMKDIEDIRSPFAEKIISFVDLKKKNLKDICYRASNESLDLLEKL 292
QY 285 LDDPSRRIPAEML-----CSPFESIIPAPHIEDLVMLPTVLRLLNVLDLDYLGNEEBY 340

DB 293 LQFNPSKRSIAENALKHKVEEFHSIIDEPTCRHIIIP----- 331
 QY 341 EDVVEEDVKECCQKGPVSLVLPKENGRCQVFEVYANMGDSKAOK 387
 DB 332 --INDNTKRVNFYRNVVYFVIMRNKFKHSNVL-----NQGSEKKEK 372

RESULT 12

DB 084VX4 PRELIMINARY; PRT; 777 AA.

AC 084VX4, 084VX4, 01-JUN-2003 (TReMBLrel. 24, Created)
 DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE A161g77720
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;

RP SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
 RA Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Huan V.W.,
 RA Ishida J., Jones T., Kamiya A., Katlin-Neumann G., Kawai J., Lam B.,
 RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
 RA Palm C.J., Quach H.T., Sakurai T., Satou M., Seki M., Southwick A.,
 RA Tang C.C., Tortum M., Wong C., Wu H.C., Yamada K., Yu G., Yuan S.,
 RA Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
 RL EMBL; BR004636; AAC02882.1; -.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR011009; Kinase like.
 DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR002290; Ser Thr kinase.
 DR InterPro; IPR008271; Ser Thr_pkin_AS.
 DR Pfam; PF00069; Pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; UNKNOWN_1.
 DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
 KW ATP-binding; Kinase; Transferase.
 SQ SEQUENCE 777 AA; 86376 MW; F2C25DA9609A878F CRC64;

Query Match 11.0%; Score 243.5; DB 2; Length 777;
 Best Local Similarity 25.9%; Pred. No. 8.2e-10;
 Matches 106; Conservative 56; Mismatches 146; Indels 101; Gaps 17;

QY 11 EPPRPLEAFGRMLVQVSRIGSGSSASVYRCCGNPSPGALKQFLPPTGTAASAAE 70
 DB 388 DPLLPKVNKGKLYQRLEKIGSGSSSEVHKVY---SSDCTYALKIKLKGRDYATA----- 440
 QY 71 YGRKRRALAEQOGSRNIVTYLGVFTIHSPNVPSCRLLLELDVSVSE----- 120
 DB 441 YGCGOIEGYLKKLKGKTNIIQL-----IDYEYTKTLLOEVLNIGTMSKDGKVEDGF 493
 QY 121 ---LLIYS-----SHOGCSMW-----MIOHCARDVLEALFLHGEYVHADLK 160
 DB 494 IYVAVLEGEIDLAMLSQKKRELEGSRTIDENWLKPYWQOIIQAVNTTIEEKIVHSDLK 553
 QY 161 PNNILWSAENECFKLIDFGLSFKENQDVYKIQT-----GYRDEALONCIAQAGIQ 214
 DB 554 PANFL--LVKGFLLKIDFGIA--KAINSDTNIQRDSQVGLTWSSEAFVCSNDENG-- 608
 QY 215 SDPECTSAVDNLSGLIILLEMFGMKLKTIVRSQEKANSALID--HIPASAVVNAAL 272
 DB 609 NTKICRPSPIWISGLCYOMVYG--RTFPADYKTFW--AKFKVITDPNHEITYNQLSNP-- 664
 QY 273 PAYHLMDLKSMLHDDPSRIRIPAEMLCSPFFSIPFAPH----- 311

DB 665 ---WLIDLMKKCLANDBRNGRWIRPELQHPFLAPPI-PHEPQVTKILFSLIAESCGSD 720
 QY 312 -----IEDLVMLPTFVLRLANLVDDDYGNSEBEVEDVVEKKEC 351
 DB 721 DRANSWISQLEQLLSNPAPLPR-----NDVIDSRDQOQLSRVSELC 763

RESULT 13

DB 094656 PRELIMINARY; PRT; 826 AA.

AC 094656, 094656, 01-FEB-1997 (TReMBLrel. 02, Created)
 DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE Mitogen-activated protein kinase 1, serine/threonine protein
 DE kinase.
 GN Name=pfmap1;
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
 NCBI_TaxID=5833;

RP SEQUENCE FROM N.A.
 RA Doering C.D., Parzy D., Langleley G., Horrocks P., Carter R.,
 RA Doering C.D., Parzy D., Langleley G., Horrocks P., Carter R.,
 RT "A map kinase homologue from the human malaria parasite Plasmodium
 RT falciparum.";
 RL Gene 177:1-6(1996).
 RN (2)

RP SEQUENCE FROM N.A.
 RC STRAIN=IT04;
 RA Doering C.,
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- Similarity: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; X82646; CAA57972.1; -.
 DR PIR; JC5153; JC5153.
 DR HSSP; Q6539; 1KV1..
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR011009; Kinase like.
 DR InterPro; IPR003527; MAP_kin.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser Thr_kinase.
 DR InterPro; IPR008271; Ser Thr_pkin_AS.
 DR Pfam; PF00069; Pkinase; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR PROSITE; PS01351; MARK; UNKNOWN_1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 826 AA; 96645 MW; A668EF77E73438A CRC64;

Query Match 11.0%; Score 243.5; DB 2; Length 826;
 Best Local Similarity 25.8%; Pred. No. 8.9e-10;
 Matches 105; Conservative 58; Mismatches 145; Indels 99; Gaps 20;

QY 23 NOVQSRIGSGSSASVYRCCGNPSPGALKQFLPPTGTAASAA--AEYFRKERAL 80
 DB 23 YDILKKVKGAGVGVFKGCKKNKIV--AVKKIF-----GAFQCTDQRTFR-ELIFL 74
 QY 81 EOLQGRNIVTYLGVFTIHSPNVPSCRLLLELDVSVSEL-----LYSSHOGCSMMMIQ 136
 DB 75 YELNGHNDIITKMDYK--AKNDNDIYIIFPMTDLHEVTKADLLEIRK-----K 124
 QY 137 HCARVDLEALFLHGEYVHADLKPRNIIWSAENEC-FGLIDFGLS-----FKEGNQDV 189
 DB 125 YIIYQLRALKYIHSGGLLRDIPKSNII--VNSECHIKVADFGIARISISTVNNKVP 182

```

QY 190 --KYIOTDGYRAPEALONCLAQIOSPTECTSAVDLSLGIILLEMFSMKL---KKT 244
DB 183 LNDYATRYRABEI-----LLGSTHYTEDVDMWSLGCINGELLGKRLFTGNST 232
QY 245 VNSQE-----WKANSSAIID--HIPASKAVVN-AAIPAYHNR-----DLIKSM 284
DB 233 MNOLEKTIIOVIGKPNKKDIEDIRSPAEKISSFVDLKKKNLKDICYKASNESLDLLEKL 292
QY 285 LHDDPSRRIPAEKAL-----CSPFESIPFAPIHEDLVMLPTPYRLNLVLDYLGNEEY 340
DB 293 LQPNPSKRISAENALKHKVEEFHSIIDEPTRHIIITP----- 331
QY 341 EDVEDVEKEECQKYGPNVSLVPEKNPGRGQVFVEYANAGSKAOK 387
DB 332 --INDNTKRVNPFYRNVVYFVIMRRNKKFHSNVL---NGESKKEEK 372

RESULT 14
Q8ILFO PRELIMINARY; PRT; 914 AA.
ID 08ILFO;
AC 08ILFO;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mitogen-activated protein kinase 1.
GN ORFNames=PF14_0294;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=63329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Carlson J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyte S.,
RA Chan M.S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Valdivia A.B.,
RA Martin D.M., Fairclamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings I.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum."
RL Nature 419:498-511(2002).
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AE014820; AAN36907.1; -.
DR HSSP; Q16539; 1KVL.
DR GO; GO:000524; P:ATP binding; IEA.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR003527; MAP kin.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_kinase.
DR Pfam; PF00069; Pkinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS01351; MAPK; UNKNOWN 1.
DR PROSITE; PS0107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
KW SEQUENCE 914 AA; 107277 MW; CAFA486A8051FA72 CRC64;

```

```

Query Match 11.0%; Score 243.5; DB 2; Length 914;
Best Local Similarity 25.8%; Pred. No. 1e-09;
Matches 105; Conservative 59; Mismatches 145; Indels 99; Gaps 20;

```

```

QY 23 MOVOSRLSGSSASVYRVRCGNGSPGALQKFLPCTGTAASA--AEYGRKRRAL 80
DB 23 YDILKKVGKAGVGVFRGCRCKKNIV--AVKKIF-----GAPQNCIDQRTTR-ELIFL 74

```

```

QY 81 EQLQGRNIVTLVGYFTTHFSPNPSRCLLELDVSVSEL---LTISSHQCSMMMIQ 136
DB 75 YELNGNDIITIKLMDYIK---AKNDNDIYLIFDFMTELDHEVIAKADLEIRK-----K 124
QY 137 HCARVTEPLALHHEGVYADLKPRNILMSAENSC-FELIDFGS-----FKENQDV 189
DB 125 YIIYQLRLKTIHSGGLHRIKPSNLI--VNSCHIKVADFGIARISITVNEKVP 182
QY 190 --KYIOTDGYRAPEALONCLAQIOSPTECTSAVDLSLGIILLEMFSMKL---KKT 244
DB 183 LNDYATRYRABEI-----LLGSTHYTEDVDMWSLGCINGELLGKRLFTGNST 232
QY 245 VNSQE-----WKANSSAIID--HIPASKAVVN-AAIPAYHNR-----DLIKSM 284
DB 293 LQPNPSKRISAENALKHKVEEFHSIIDEPTRHIIITP----- 331
QY 341 EDVEDVEKEECQKYGPNVSLVPEKNPGRGQVFVEYANAGSKAOK 387
DB 332 --INDNTKRVNPFYRNVVYFVIMRRNKKFHSNVL---NGESKKEEK 372

RESULT 15
Q8AYR3 PRELIMINARY; PRT; 395 AA.
ID 08AYR3;
AC 08AYR3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Calcium/calmodulin-dependent protein kinase (Cam-KI protein).
GN Name=Cam-KI;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Saneyoshi T., Kume S., Mikoshiba K.;
RT "Calcium/calmodulin-dependent protein kinase I in Xenopus laevis."
RL Comp. Biochem. Physiol. B, Comp. Biochem. 134:499-507(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Oocytes;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.U., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mollath S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.U., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallajon D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,
RA Faley J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C.,
RA Rodriguez A.C., Gilmwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallus D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Oocytes;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,

```

RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative";
RL Dev. Dyn. 225:384-391(2002).
RN (4)
RP SEQUENCE FROM N.A.
RC TISSUE=Oocytes;
RA Klein S., Strausberg R.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AB082399; BAC13847.1; -;
DR EMBL; BC070745; AAH70745.1; -;
DR HSSP; Q63450; 1A06.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_kin_as.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 395 AA; 43801 MW; 5381D04890B5D3AE CRC64;

Query Match 11.0%; Score 242.5; DB 2; Length 395;
Best Local Similarity 27.7%; Pred. No. 4.2e-10;
Matches 91; Conservative 47; Mismatches 108; Indels 83; Gaps 16;

QY 4 SGCAMGAEPPEPRLEAFGRIMQVQSRIGSSGSASV-----YVRCCGNGSPGPA 52
DB 9 SGTGCKKQ---VEDIKRMQFKFVGTGAFSEVLAEEKATGKLFVVKCI-----PKKA 59
QY 53 LKQFLPFGTTGAASAAYGFRKERAALBQLQGRNIVTLVGVFTIHFSNVPSRCLL- 111
DB 60 LK-----GKESIENEIVLKKIK-HENIVALEDIY-----ESPSHLVLM 99
QY 112 -----ELLDVSVSRLLVSSHQCSMMMIQHCARDVLEALAFIHGGYVHADLKPNTL 165
DB 100 QLVSGGELPD-RIVEKGFTEKXASTL-----IRQVLDAVSYLHRLGIVHRDLKPENL 152
QY 166 WSAENECEFKLI--DFGLSFKEGNQDYKYLQ--TDGYRAPPAELQNCIAQAGLSDECTG 221
DB 153 YFSQVESKIMISDFGLSKMEGKGDVWSTACGTPGVAPEV-----LAQ-----KPYSK 201
QY 222 AYDLMLSLGIIILEMFGMKLKTIVRSQEWKANSALIIDHIFASKAVVNAALPAY-----H 276
DB 202 AYDCWISIGVAYIYLLCGY-----PPFYDENSKLFEQTL--KADYEFDPFMDISDS 252
QY 277 LRDLSKMLHDPSPRRIPAEMLCSPPFS 305
DB 253 AKDFIQNLMKEDPNKRYACGQALRHPWIA 281

Search completed: November 29, 2004, 11:46:52
Job time : 126.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2004, 11:26:47 ; Search time 109.5 Seconds
(without alignments)
1372.672 Million cell updates/sec

Title: US-10-798-532-2

Perfect score: 2206
Sequence: 1 MAGSCAMAEPRPRLFARG.....VATFVPLSAVKRGVLTLL 419

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

1: geneseqp1980s:*\n2: geneseqp1980s:*\n3: geneseqp2000s:*\n4: geneseqp2001s:*\n5: geneseqp2002s:*\n6: geneseqp2003as:*\n7: geneseqp2003bs:*\n8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2198	99.6	419	4	AAG65767 Human pro
2	2198	99.6	419	5	ABP64797 Human pro
3	2191	99.3	419	4	AAU03534 Human pro
4	2191	99.3	419	4	AAE11769 Human kin
5	2190.5	99.3	418	3	AAV70305 Human tra
6	2187.5	99.2	418	3	AAV70306 Human tra
7	2186	99.1	419	7	ABPD48093 Rat Prote
8	483.5	21.9	137	5	ABP64662 Human ORF
9	243.5	11.0	638	3	AAG38502 Arabidops
10	243.5	11.0	720	3	AAG38501 Arabidops
11	243.5	11.0	777	3	AAG38500 Arabidops
12	239.5	10.9	544	2	AAW85028 CDK2-gree
13	238	10.8	298	7	ADD21398 Golden ha
14	236	10.7	435	6	AAE31916 Yeast Hog
15	235	10.7	367	3	AAE22137 Rat Ectres
16	235	10.7	435	6	AAE31917 Yeast Hog
17	235	10.7	435	6	AAE31922 Yeast Hog
18	235	10.7	435	6	AAE31921 Yeast Hog
19	235	10.7	435	6	AAE31914 Yeast Hog
20	235	10.7	435	2	AAE31913 Yeast Hog
21	235	10.7	544	2	AAW85029 Green flu
22	234.5	10.6	433	3	AAE19094 Protein e
23	234.5	10.6	433	8	ADP87472 S cerevis
24	234	10.6	298	2	AAW95689 Homo sapi
25	234	10.6	298	7	ADD21402 Chinese h

26	234	10.6	298	7	ADD21406	ADD21406 Rat cycli
27	234	10.6	298	2	ADD21394	ADD21394 Mouse cyc
28	234	10.6	416	7	AAW06795	AAW06795 High Osmo
29	234	10.6	435	6	AAE31910	AAE31910 Yeast Hog
30	234	10.6	435	6	AAE31915	AAE31915 Yeast Hog
31	234	10.6	435	6	AAE31918	AAE31918 Yeast Hog
32	234	10.6	435	6	AAE31919	AAE31919 Yeast Hog
33	234	10.6	435	6	AAE31926	AAE31926 Yeast Hog
34	234	10.6	435	6	AAE31911	AAE31911 Yeast Hog
35	234	10.6	435	6	AAE31909	AAE31909 Yeast Hog
36	234	10.6	435	6	AAE31912	AAE31912 Yeast Hog
37	234	10.6	435	6	AAE31927	AAE31927 Yeast Hog
38	234	10.6	298	6	ABR58686	ABR58686 Human can
39	233	10.6	298	7	ADD21410	ADD21410 Goldfish
40	233	10.6	298	7	ADD21392	ADD21392 Human cyc
41	233	10.6	298	7	ADP45038	ADP45038 Human kin
42	233	10.6	298	8	AD157232	AD157232 NK1AWRE d
43	233	10.6	298	8	ADJ66599	ADJ66599 cdk2 prot
44	233	10.6	392	4	ABB63118	ABB63118 Drosophila
45	232.5	10.5	355	4	AAE50055	AAE50055 Murine De

ALIGNMENTS

RESULT 1
ID AAG65767 standard; protein; 419 AA.

AC AAG65767;

DT 07-JUN-2002 (first entry)

DE Human protein kinase polypeptide 13887.

XX Protein kinase; 3714; 16742; 23546; 13887; cancer; bone disorder; human;

KW cytostatic; antiinflammatory; immunosuppressive; cardiac; hepatocytic;

KM Gene therapy.

OS Homo sapiens.

PN WO200173050-A2.

PD 04-OCT-2001.

PF 23-MAR-2001; 2001WO-US009483.

PR 24-MAR-2000; 2000US-0191846P.

PA (MILL-) MILLENNIUM PHARM INC.

PI Meyers R;

DR WPI; 2001-611632/70.

DR N-PSDB; AA166828, AA166829.

PT New human protein kinase polypeptides; 3714, 16742, 23546 and 13887, useful in diagnosis of cancer or cellular proliferation or to treat such conditions.

PS Claim 4; Fig 10A-C; 16pp; English.

XX The invention provides novel human protein kinase polypeptides, 3714, 16742, 23546 and 13887 and nucleic acid molecules encoding them. The protein kinase polypeptides can be expressed by standard recombinant methodology. 3714, 16742, 23546 or 13887 nucleic acids and polypeptides are useful for diagnostic and screening methods to identify subjects (at risk of) having cancer or cellular proliferation and/or differentiation disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and modulators are useful for the treatment of cancer, particularly colon cancer or cellular proliferation and/or differentiation disorders. Other disorders associated with 3714, 16742, 23546 or 13887 expression or

CC activity that can be treated include bone related disorders, inflammatory
 CC disorders, autoimmune diseases, cardiovascular disorders and liver
 CC diseases. The present sequence represents a human protein kinase
 CC polypeptide 13887

XX Sequence 419 AA;

Query Match 99.6%; Score 2198; DB 4; Length 419;
 Best Local Similarity 99.8%; Pred. No. 2.1e-229;
 Matches 418; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAGSGCAGMBPFPFLFAGRLMVOVSRLGSSGSASVYRVRCGNGSPGALKQFLPPG 60
 DB 1 MAGSGCAGMBPFPFLFAGRLMVOVSRLGSSGSASVYRVRCGNGSPGALKQFLPPG 60
 QY 61 TTGAASAASAEYGRKRRALAEQLQGHNRNVTLYGVFTTHFSNVPSRCLLELDVSVSE 120
 DB 61 TTGAASAASAEYGRKRRALAEQLQGHNRNVTLYGVFTTHFSNVPSRCLLELDVSVSE 120
 QY 121 LLLYSSHQGCSMMWIOHCARDVLEALAFHHEGYHADLKPRNIIIMSANECEFKLIDFGL 180
 DB 121 LLLYSSHQGCSMMWIOHCARDVLEALAFHHEGYHADLKPRNIIIMSANECEFKLIDFGL 180
 QY 181 SFKEGNQDVYKTIQTDGYRAPEALQNCIAQAGLQSDTECTSAVDLMSLGIIILEMFSGMK 240
 DB 181 SFKEGNQDVYKTIQTDGYRAPEALQNCIAQAGLQSDTECTSAVDLMSLGIIILEMFSGMK 240
 QY 241 LKHTVRSQEWKANSNIIIDHIFASKAVVNAAPAYHLDLIXSMHDDPSRRIIPAEMLC 300
 DB 241 LKHTVRSQEWKANSNIIIDHIFASKAVVNAAPAYHLDLIXSMHDDPSRRIIPAEMLC 300
 QY 301 SPFFSIPFAPHIEDLVMLPTPVRLINVLDDDYLGNEBEYEDVDVEKCECKYGPVSL 360
 DB 301 SPFFSIPFAPHIEDLVMLPTPVRLINVLDDDYLGNEBEYEDVDVEKCECKYGPVSL 360
 QY 361 LVPKENPGRGVVEVYANAGDSKAAQKLLTGMFDGKFVATFYPLSAVKRGYLYQTLL 419
 DB 361 LVPKENPGRGVVEVYANAGDSKAAQKLLTGMFDGKFVATFYPLSAVKRGYLYQTLL 419

RESULT 2
 ABB64797
 ID ABB64797 standard; protein; 419 AA.

XX ABB64797;
 XX
 DT 25-FEB-2003 (first entry)
 XX
 DE Human protein SEQ ID 457.
 XX

Human, expressed sequence tag; EST; hematopoietic disorder;
 central nervous system disease; viral infection;
 peripheral nervous system disease; non-healing wound; infectious disease;
 immune deficiency; immune disorder; bacterial infection; allergy; cancer;
 fungal infection; autoimmune disorder; coagulation disorder; neutrotic;
 antiallergic; antiinflammatory; immunosuppressive; neuroprotective;
 cytostatic; haemostatic; virucide; antibacterial; fungicide;
 immunostimulant; cerebroprotective.

OS Homo sapiens.
 PN WO200259260-A2.
 PD 01-AUG-2002.
 XX
 XX 16-NOV-2001; 2001MO-US042950.
 PR 17-NOV-2000; 2000US-00714936.
 XX
 XX (HYSE-) HYSEQ INC.
 PA
 XX Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
 PI Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;

XX WPI; 2002-590824/63.
 DR N-PSDB; ABQ9383.
 XX

PT New isolated polynucleotide, useful in research, diagnostic or
 PT therapeutic methods, e.g. preventing or treating disorders involving
 PT aberrant protein expression or biological activity.
 XX

PS Claim 20; SEQ ID NO 457; 394bp; English.

CC The present invention relates to novel human coding sequences (ABQ99268-
 CC ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in
 CC therapeutic, diagnostic and research methods. The polynucleotides may be
 CC used in the field of molecular biology as hybridisation probes, primers
 CC for PCR, for chromosome and gene mapping, for the recombinant production
 CC of protein, or in generation of anti-sense DNA or RNA. The
 CC polynucleotides are useful in diagnostics as expressed sequence tags
 CC (ESTs) for identifying expressed genes or for physical mapping of the
 CC human genome. The proteins may be used as molecular weight markers, or as
 CC nutritional sources or supplements. The proteins may be used to maintain
 CC and expand cell population in a totipotent or pluripotent state
 CC useful for re-engineering damaged or diseased tissues, transplantation,
 CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The
 CC polynucleotides and proteins are useful for preventing, treating or
 CC ameliorating disorders involving aberrant protein expression or
 CC biological activity, e.g. hematopoietic disorders, central/peripheral
 CC nervous system diseases, mechanical and traumatic disorders, non-healing
 CC wounds, immune deficiencies and disorders, infectious diseases caused by
 CC viral, bacterial or fungal infection, autoimmune disorders, allergic
 CC reactions and conditions, coagulation disorders, or cancer. The
 CC polynucleotide sequences of the invention were assembled from ESTs
 CC isolated mainly by sequencing by hybridisation, and in some cases,
 CC sequences obtained from one or more public databases. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPRO at
 CC ftp.wipro.int/pub/published_pct_sequences
 XX

XX Sequence 419 AA;

Query Match 99.6%; Score 2198; DB 5; Length 419;
 Best Local Similarity 99.8%; Pred. No. 2.1e-229;
 Matches 418; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAGSGCAGMBPFPFLFAGRLMVOVSRLGSSGSASVYRVRCGNGSPGALKQFLPPG 60
 DB 1 MAGSGCAGMBPFPFLFAGRLMVOVSRLGSSGSASVYRVRCGNGSPGALKQFLPPG 60
 QY 61 TTGAASAASAEYGRKRRALAEQLQGHNRNVTLYGVFTTHFSNVPSRCLLELDVSVSE 120
 DB 61 TTGAASAASAEYGRKRRALAEQLQGHNRNVTLYGVFTTHFSNVPSRCLLELDVSVSE 120
 QY 121 LLLYSSHQGCSMMWIOHCARDVLEALAFHHEGYHADLKPRNIIIMSANECEFKLIDFGL 180
 DB 121 LLLYSSHQGCSMMWIOHCARDVLEALAFHHEGYHADLKPRNIIIMSANECEFKLIDFGL 180
 QY 181 SFKEGNQDVYKTIQTDGYRAPEALQNCIAQAGLQSDTECTSAVDLMSLGIIILEMFSGMK 240
 DB 181 SFKEGNQDVYKTIQTDGYRAPEALQNCIAQAGLQSDTECTSAVDLMSLGIIILEMFSGMK 240
 QY 241 LKHTVRSQEWKANSNIIIDHIFASKAVVNAAPAYHLDLIXSMHDDPSRRIIPAEMLC 300
 DB 241 LKHTVRSQEWKANSNIIIDHIFASKAVVNAAPAYHLDLIXSMHDDPSRRIIPAEMLC 300
 QY 301 SPFFSIPFAPHIEDLVMLPTPVRLINVLDDDYLGNEBEYEDVDVEKCECKYGPVSL 360
 DB 301 SPFFSIPFAPHIEDLVMLPTPVRLINVLDDDYLGNEBEYEDVDVEKCECKYGPVSL 360
 QY 361 LVPKENPGRGVVEVYANAGDSKAAQKLLTGMFDGKFVATFYPLSAVKRGYLYQTLL 419
 DB 361 LVPKENPGRGVVEVYANAGDSKAAQKLLTGMFDGKFVATFYPLSAVKRGYLYQTLL 419

RESULT 3

AAU03534
 ID AAU03534 standard; protein; 419 AA.
 AC AAU03534;
 XX
 DT 12-SEP-2001 (first entry)
 XX
 DE Human protein kinase #34.
 XX
 KW Human; protein kinase; PK; STK; cancer; cardiovascular disease;
 KW metabolic disorder; immune related disease; neurological disorder;
 KW neurodegenerative disorder; inflammatory disorder; infectious disease;
 KW reproductive disorder.
 OS Homo sapiens.
 PN MO200138503-A2.
 PD 31-MAY-2001.
 PF 22-NOV-2000; 2000MO-US032085.
 PR 24-NOV-1999; 99US-0167482P.
 XX
 PA (SUGB-) SUGEN INC.
 PI Piowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
 PI Flanagan P, Clardy D;
 XX
 DR WPI; 2001-343950/36.
 XX N-PSDB; AAS06734.
 PT Nucleic acids encoding human kinase polypeptides, useful for preventing
 PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
 PT neuronal-associated diseases, and microbial infections.
 XX
 PS Claim 7; Fig 2; 433pp; English.
 XX
 CC AAU03501-AAU03557 represent novel human protein kinases #1-57. The novel
 CC protein kinases have been identified as members of the tyrosine or
 CC serine/threonine kinase (PTK and STK) families. The polynucleotides
 CC encoding protein kinases and the polypeptides may be used in the
 CC prevention, diagnosis and treatment of diseases associated with
 CC inappropriate kinase expression. For example, they may be used to treat
 CC cancers (especially cancers of haematopoietic origin), cardiovascular
 CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
 CC immune related diseases (e.g. rheumatoid arthritis), neurological
 CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
 CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
 CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
 CC Additionally, polynucleotides encoding protein kinases may be used for
 CC gene therapy and as DNA probes in diagnostic assays. The protein kinase
 CC polypeptides may be used as antigens in the production of antibodies
 CC against the protein kinases and in assays to identify modulators of
 CC protein kinase expression and activity
 CC
 SQ Sequence 419 AA;
 Query Match 99.3%; Score 2191; DB 4; Length 419;
 Best Local Similarity 99.5%; Pred. No. 1.2e-228;
 Matches 417; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 181 SPKSGNODVKYIQTGGRAPAEALONCLAQGLQSTECTSAVDLMSLGIILLBMPSGMK 240
 DB |||||||
 181 SPKSGNODVKYIQTGGRAPAEALONCLAQGLQSTECTSAVDLMSLGIILLBMPSGMK 240
 QY 241 LKHTVASQEMKANSNAIIDHIFASKAVVNAAPAYHLRLIKSMHDDPSRRIPAEMLC 300
 DB |||||||
 241 LKHTVASQEMKANSNAIIDHIFASKAVVNAAPAYHLRLIKSMHDDPSRRIPAEMLC 300
 QY 301 SPFFSIPFAPHIEDLVMLPTPLRLINLVDDYLGNEEYEDVEDYKECCQYGPVSL 360
 DB |||||||
 301 SPFFSIPFAPHIEDLVMLPTPLRLINLVDDYLGNEEYEDVEDYKECCQYGPVSL 360
 QY 361 LVPKRNPRGQVFEVYANAGDSKAOKLITGRMPGKPVATFPYLSAYKRGVLYOTLL 419
 DB |||||||
 361 LVPKRNPRGQVFEVYANAGDSKAOKLITGRMPGKPVATFPYLSAYKRGVLYOTLL 419
 RESULT 4
 ID AAEL1769 standard; protein; 419 AA.
 AC AAEL1769;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human kinase (PKIN) -3 protein.
 XX
 KW Human kinase; PKIN; gene therapy; adenocarcinoma; immune disorder; gout;
 KW cancer; allergy; sarcoma; leukaemia; acquired immune deficiency syndrome;
 KW AIDS; Addison's disease; microbial infection; inflammation; osteoporosis;
 KW atherosclerosis; cardiovascular disease; myocardial infarction; anaemia;
 KW myasthenia gravis; cataracts; cataract; growth and development disorder;
 KW seizure disorder; pulmonary embolism; Gaucher's disease; lipid disorder;
 KW lipid storage disease; Pick's disease; Tay-Sachs disease; renal disease;
 KW asthma; obesity; restorative therapy; cytostatic; immunomodulatory;
 KW antimicrobial; cardiovascular; antiinflammatory; vaccine.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 23..304
 FT /note= "Eukaryotic protein kinase domain"
 FT 48..289
 FT /note= "Protein kinase domain"
 FT 74..293
 FT /note= "Protein kinase domain"
 FT 108..241
 FT /note= "Protein kinase domain"
 FT 148..166
 FT /note= "Tyrosine kinase catalytic domain"
 FT 223..245
 FT /note= "Tyrosine kinase catalytic domain"
 FT 223..241
 FT /label= "Transmembrane_domain"
 FT 273..295
 FT /note= "Tyrosine kinase catalytic domain"
 FT 305..344
 FT /note= "Serine/threonine protein kinase"
 FT 319..404
 FT /note= "Ribonucleoprotein repeat"
 FT 320..412
 FT /note= "Splicing factor like protein"
 FT Region
 PN MO200181555-A2.
 PD 01-NOV-2001.
 XX
 XX 20-APR-2001; 2001MO-US012992.
 PF
 XX
 XX 20-APR-2000; 2000US-0199021P.
 PR 28-APR-2000; 2000US-0200226P.
 PR 05-MAY-2000; 2000US-0202339P.

PR 11-MAY-2000; 2000US-0203505P.
 PR 18-MAY-2000; 2000US-0205564P.
 PR 26-MAY-2000; 2000US-0207739P.
 PR 01-JUN-2000; 2000US-0208795P.
 XX
 XX (INCYTE) INCYTE GENOMICS INC.
 XX
 PI Yue H, Gandhi AR, Tribouley CM, Kearney L, Griffin JA, Nguyen DB;
 PI Bandman O, Lu DM, Lal P, Burford N, Khan FA, Walla NK, Yao MG;
 PI Patterson C, Burrill JD, Marcus GA, Zingler KA, Reardon SA, Lu Y;
 PI Policky JL, Thornton M, Tang YT, Hafalia A, Elliott VS, Baughn MR;
 PI Walsh RT, Ramkumar J, Borowsky ML, Au-Young J, Hillman JL,
 PI Gururajan R;
 XX
 DR WPI: 2001-611740/70.
 DR N-PSDB; AAD18618.
 XX
 PT Human kinases and nucleic acids, useful for preventing diagnosing and
 PT treating cancers, inflammation and immune disorders.
 XX
 PS Claim 1; Page 125-126; 16pp; English.
 XX
 CC The present invention relates to human kinases (PKIN) and the nucleic
 CC acids encoding them. PKIN is used as vaccine and in gene therapy. PKIN is
 CC used in the prevention, diagnosis and treatment of diseases cancers,
 CC adenocarcinoma, leukaemia, sarcoma, immune disorder, Addison's disease,
 CC acquired immune deficiency syndrome (AIDS), anaemia, asthma, allergies,
 CC gout, microbial infections, cardiovascular disease and/or inflammation,
 CC myasthenia gravis, atherosclerosis, cirrhosis, osteoporosis, myocardial
 CC infarction, cataract, growth and development disorder, seizure disorder,
 CC pulmonary embolism, Gaucher's disease, lipid disorder, lipid storage
 CC disease, Pick's disease, Tay-Sachs disease, renal disease and obesity.
 CC PKIN may be used to treat disorders associated with decreased PKIN
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of PKIN by expressing inactive proteins or to
 CC supplement the patient's own production of PKIN. PKIN nucleic acids may be
 CC used to produce the PKIN polypeptide, by inserting the nucleic acids into
 CC a host cell and culturing the cell to express the protein. PKIN nucleic
 CC acid and its complementary sequences may also be used as DNA probes in
 CC diagnostic assays to detect and quantitate the presence of similar
 CC nucleic acid sequences in samples and therefore which patients may be in
 CC need of restorative therapy. The present sequence is human PKIN-3 protein
 CC
 XX
 SQ Sequence 419 AA;
 Query Match 99.3%; Score 2191; DB 4; Length 419;
 Best Local Similarity 99.5%; Pred. No. 1.2e-228;
 Matches 417; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 XX
 QY 1 MAGSGCAGCAEPPRFLAEFGRLMOWOSRLGSSGSSASVYRVRCGNGSPGALKQFLPPG 60
 DB 1 MAGSGCAGCAEPPRFLAEFGRLMOWOSRLGSSGSSASVYRVRCGNGSPGALKQFLPPG 60
 QY 61 TTGAASAABEYGRKERRALEOLQGRNIVTLVYGVFTTHSPNVRSCILLLELVVSE 120
 DB 61 TTGAASAABEYGRKERRALEOLQGRNIVTLVYGVFTTHSPNVRSCILLLELVVSE 120
 QY 61 TTGAASAABEYGRKERRALEOLQGRNIVTLVYGVFTTHSPNVRSCILLLELVVSE 120
 DB 61 TTGAASAABEYGRKERRALEOLQGRNIVTLVYGVFTTHSPNVRSCILLLELVVSE 120
 QY 121 LLLYSSHOGCSMMWICHACARDVLEALAFHLHBEYVHADLPKRIILMSAENECKLIDFGL 180
 DB 121 LLLYSSHOGCSMMWICHACARDVLEALAFHLHBEYVHADLPKRIILMSAENECKLIDFGL 180
 QY 121 LLLYSSHOGCSMMWICHACARDVLEALAFHLHBEYVHADLPKRIILMSAENECKLIDFGL 180
 DB 121 LLLYSSHOGCSMMWICHACARDVLEALAFHLHBEYVHADLPKRIILMSAENECKLIDFGL 180
 QY 181 SFKEGNDVYIQTDDYRAPEALQNLQAGLQSTDECTSAVDLMSLGIILLEMFSGMK 240
 DB 181 SFKEGNDVYIQTDDYRAPEALQNLQAGLQSTDECTSAVDLMSLGIILLEMFSGMK 240
 QY 241 LKHTVRSQEWKANSIIDIHIFASKAVVNAIPAYHLRDLIKSMHDDPSRRIPAMALC 300
 DB 241 LKHTVRSQEWKANSIIDIHIFASKAVVNAIPAYHLRDLIKSMHDDPSRRIPAMALC 300
 QY 301 SPFFSIPPAHTIEDVLMPTPVRLNLVLDGDEEEDVVEVKECKQYGVVSL 360
 DB 301 SPFFSIPPAHTIEDVLMPTPVRLNLVLDGDEEEDVVEVKECKQYGVVSL 360

QY 361 LVPKPNPGRGQVFEVYANAGDSKAAQKLTGRMFDGKFAVATFPLSAVKRGYLTLL 419
 DB 361 LVPKPNPGRGQVFEVYANAGDSKAAQKLTGRMFDGKFAVATFPLSAVKRGYLTLL 419
 XX
 XX RESULT 5
 XX AAY70305
 XX ID AAY70305 standard; protein; 418 AA.
 XX
 XX AAY70305;
 XX
 XX 06-JUN-2000 (first entry)
 XX
 XX Human wild type serine/threonine kinase KIS (hKIS) protein.
 XX
 XX KIS: human; hKIS; serine/threonine kinase; cell proliferation; G1 phase;
 XX cyclin-dependent kinase inhibitor; CKI; p27; modulator; treatment;
 XX cell proliferative disease; vascular disorder; gene therapy; restenosis;
 XX atherosclerosis.
 XX
 XX Homo sapiens.
 XX
 XX Key Location/Qualifiers
 XX Misc-difference 186
 XX FT /note= "Encoded by AATCAG"
 XX FT
 XX
 XX MO20001165-A1.
 XX
 XX 02-MAR-2000.
 XX
 XX 20-AUG-1999; 99WO-US018903.
 XX
 XX 21-AUG-1998; 98US-0097710P.
 XX
 XX (NABE/) NABEL G J.
 XX (NABE/) NABEL E G.
 XX
 XX Nabel GJ, Nabel EG;
 XX
 XX WPI: 2000-237648/20.
 XX N-PSDB; AA251355.
 XX
 XX Novel serine/threonine kinase hKIS polynucleotides and polypeptides used
 XX PT for inhibiting the cyclin kinase inhibitor p27, and so alter cell
 XX PT proliferation.
 XX
 XX Example 1; Page 57-58; 70pp; English.
 XX
 XX The present protein sequence is the wild type human KIS (hKIS), isolated
 XX CC from a human B-cell library, using a yeast two hybrid screening system.
 XX CC hKIS is a serine/threonine kinase, that acts as an inhibitory kinase of
 XX CC cyclin-dependent kinase inhibitor (CKI), p27. hKIS controls cell
 XX CC proliferation and is localised predominantly in the nucleus. It binds to
 XX CC CKI p27 and inhibits its ability to arrest cells in G1 phase. It has an
 XX CC amino terminal serine/threonine kinase consensus region and a carboxy
 XX CC terminus with 42% sequence similarity to hU2AF65, a subunit of the
 XX CC splicing factor U2AF. The hKIS sequences are used to modulate cell
 XX CC proliferation and treat cell proliferative and vascular diseases. The DNA
 XX CC sequence may be used in gene therapy to treat vascular disorders such as
 XX CC restenosis or atherosclerosis
 XX
 XX
 XX Sequence 418 AA;
 Query Match 99.3%; Score 2190.5; DB 3; Length 418;
 Best Local Similarity 99.8%; Pred. No. 1.4e-228;
 Matches 418; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 XX
 QY 1 MAGSGCAGCAEPPRFLAEFGRLMOWOSRLGSSGSSASVYRVRCGNGSPGALKQFLPPG 60
 DB 1 MAGSGCAGCAEPPRFLAEFGRLMOWOSRLGSSGSSASVYRVRCGNGSPGALKQFLPPG 60
 QY 61 TTGAASAABEYGRKERRALEOLQGRNIVTLVYGVFTTHSPNVRSCILLLELVVSE 120
 DB 61 TTGAASAABEYGRKERRALEOLQGRNIVTLVYGVFTTHSPNVRSCILLLELVVSE 120

```

Db      TTGAASAABEYGRKERALBOLQGHRIIVTLGVFTIHFSPNVPSCRLLLELDVSVSE 120
Qy      121 LLLYSHOCCSMMWMOHCARDVLEALAFIHGEGYHADLKPRNIIIMSAENECFKLIDFGL 180
Db      121 LLLYSHOCCSMMWMOHCARDVLEALAFIHGEGYHADLKPRNIIIMSAENECFKLIDFGL 180
Qy      181 SPFEGNODVKYIQTODGYRPAPEALONCLAOAGLOSTECTSAVDLSLGIILLEMFSGMK 240
Db      181 SPFEGN-DVKYIQTODGYRPAPEALONCLAOAGLOSTECTSAVDLSLGIILLEMFSGMK 239
Qy      241 LKHTVRSQEWKANSASAIIDHIFASKAVVNAIPAVALRDILIKSMHDDPSRRIIPAEMLC 300
Db      240 LKHTVRSQEWKANSASAIIDHIFASKAVVNAIPAVALRDILIKSMHDDPSRRIIPAEMLC 299
Qy      301 SPFFSIPFAPHIEDLVMLPTPVRLINLVDDDYLGNEEYEDVDVVEECCQYKGPVSVL 360
Db      300 SPFFSIPFAPHIEDLVMLPTPVRLINLVDDDYLGNEEYEDVDVVEECCQYKGPVSVL 359
Qy      361 LVPKPNRGGGVFEVYANAGDSKAAOKLLTGRMPDGKPVVATFYPASAYKRGYLYOTLL 419
Db      360 LVPKPNRGGGVFEVYANAGDSKAAOKLLTGRMPDGKPVVATFYPASAYKRGYLYOTLL 418

RESULT 6
AA70306
ID      AA70306 standard; protein; 418 AA.
AC      AA70306;
XX      06-JUN-2000 (first entry)
DE      Human transdominant mutant serine/chreonine kinase KIS (hKIS) protein.
XX      KIS; human; hKIS; serine/chreonine kinase; transdominant mutant; p27;
XX      G1 phase; cell proliferation; cyclin-dependent kinase inhibitor; CKI;
XX      modulator; treatment; cell proliferative disease; vascular disorder;
XX      gene therapy; atherosclerosis; restenosis.
OS      Homo sapiens.
XX      Key Location/Qualifiers
FH      Misc-difference 54
FT      /note= "Wild type Lys substituted with Arg"
FT      Misc-difference 186
FT      /note= "Encoded by AATCAG"
XX      WO20001165-A1.
XX      02-MAR-2000.
XX      20-AUG-1999; 99WO-US018903.
XX      21-AUG-1998; 98US-0097710P.
XX      (NABE/) NABEL G J.
XX      (NABE/) NABEL E G.
XX      Nabel GJ, Nabel EG;
XX      WPI; 2000-237648/20.
XX      N-PSDB; AA251356.
XX      Novel serine/threonine kinase hKIS polynucleotides and polypeptides used
XX      for inhibiting the cyclin kinase inhibitor p27, and so alter cell
XX      proliferation.
XX      Example 1; Page 59-61; 70pp; English.
XX      The present protein sequence is the transdominant mutant human KIS
XX      (hKIS), constructed by site directed mutagenesis. A single nucleotide
XX      substitution (A to G) results in a lysine to arginine change in the
XX      protein sequence. hKIS is a serine/chreonine kinase, that acts as an
XX      inhibitory kinase of cyclin-dependent kinase inhibitor (CKI), p27. hKIS

```

```

CC      controls cell proliferation and is localised predominantly in the
CC      nucleus. It binds to CKI p27 and inhibits its ability to arrest cells in
CC      G1 phase. The hKIS sequences are used to modulate cell proliferation and
CC      treat cell proliferative and vascular diseases. The polynucleotide
CC      sequence may be used in gene therapy to treat vascular disorders such as
CC      restenosis or atherosclerosis
XX      Sequence 418 AA:
SQ
Query Match      99.2%; Score 2187.5; DB 3; Length 418;
Best Local Similarity 99.5%; Pred. No. 2.9e-228;
Matches 417; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
Qy      1 MAGSGCAGABEPRLLEAFGRIMQVSRUGSSSSAYVRVRCGPGSPGALXQFLPPG 60
Db      1 MAGSGCAGABEPRLLEAFGRIMQVSRUGSSSSAYVRVRCGPGSPGALXQFLPPG 60
Qy      61 TTGAASAABEYGRKERALBOLQGHRIIVTLGVFTIHFSPNVPSCRLLLELDVSVSE 120
Db      61 TTGAASAABEYGRKERALBOLQGHRIIVTLGVFTIHFSPNVPSCRLLLELDVSVSE 120
Qy      121 LLLYSHOCCSMMWMOHCARDVLEALAFIHGEGYHADLKPRNIIIMSAENECFKLIDFGL 180
Db      121 LLLYSHOCCSMMWMOHCARDVLEALAFIHGEGYHADLKPRNIIIMSAENECFKLIDFGL 180
Qy      181 SPFEGNODVKYIQTODGYRPAPEALONCLAOAGLOSTECTSAVDLSLGIILLEMFSGMK 240
Db      181 SPFEGN-DVKYIQTODGYRPAPEALONCLAOAGLOSTECTSAVDLSLGIILLEMFSGMK 239
Qy      241 LKHTVRSQEWKANSASAIIDHIFASKAVVNAIPAVALRDILIKSMHDDPSRRIIPAEMLC 300
Db      240 LKHTVRSQEWKANSASAIIDHIFASKAVVNAIPAVALRDILIKSMHDDPSRRIIPAEMLC 299
Qy      301 SPFFSIPFAPHIEDLVMLPTPVRLINLVDDDYLGNEEYEDVDVVEECCQYKGPVSVL 360
Db      300 SPFFSIPFAPHIEDLVMLPTPVRLINLVDDDYLGNEEYEDVDVVEECCQYKGPVSVL 359
Qy      361 LVPKPNRGGGVFEVYANAGDSKAAOKLLTGRMPDGKPVVATFYPASAYKRGYLYOTLL 419
Db      360 LVPKPNRGGGVFEVYANAGDSKAAOKLLTGRMPDGKPVVATFYPASAYKRGYLYOTLL 418

RESULT 7
ADD48093
ID      ADD48093 standard; protein; 419 AA.
AC      ADD48093;
XX      29-JAN-2004 (first entry)
DE      Rat Protein AAC53031, SEQ ID NO 13791.
XX      Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX      chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX      Rattus norvegicus.
XX      WO2003016475-A2.
XX      27-FEB-2003.
XX      14-AUG-2002; 2002WO-US025765.
XX      14-AUG-2001; 2001US-0312147P.
XX      01-NOV-2001; 2001US-0346382P.
XX      26-NOV-2001; 2001US-0333347P.
XX      (GENO ) GEN HOSPITAL CORP.
XX      (FARB ) BAYER AG.
XX      Woolf C, D'urbo D, Befort K, Costigan M;
XX      WPI; 2003-268312/26.

```

DR GENBANK; AAC53031.

XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 419 AA:

Query Match 99.1%; Score 2186; DB 7; Length 419;
Best Local Similarity 99.0%; Pred. No. 4,3e-228;
Matches 415; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MGSSGGAMKAEPRPFLEAFGRLLMOWOSRIGSSSSAVYRRCGNGSPGALKQRLPG 60
DB 1 MGSSGGAMKAEPRPFLEAFGRLLMOWOSRIGSSSSAVYRRCGNGSPGALKQRLPG 60
QY 61 TTGAASAAYEYGRKERALEQOGHRNIVTLVGVFTIHPSPVPSRCLLELDVSVS 120
DB 61 TTGAASAAYEYGRKERALEQOGHRNIVTLVGVFTIHPSPVPSRCLLELDVSVS 120
QY 121 LLYSSHQCGSMWMIQHCARDVLEALAFHHEGYVHADLPKRNILWSAENECFKLIDFG 180
DB 121 LLYSSHQCGSMWMIQHCARDVLEALAFHHEGYVHADLPKRNILWSAENECFKLIDFG 180
QY 181 SPREGNOVKYITODGYRAPEALONCLAQAGQSTECTSANDLWSLIGITLLEMSGMK 240
DB 181 SPREGNOVKYITODGYRAPEALONCLAQAGQSTECTSANDLWSLIGITLLEMSGMK 240
QY 241 LKHTVASQEMKANSAAIHIFASKAVVNAALPAVYHRLDKLSMLDDPSRRIAPAMALC 300
DB 241 LKHTVASQEMKANSAAIHIFASKAVVNAALPAVYHRLDKLSMLDDPSRRIAPAMALC 300
QY 301 SPFFSIPFAPHIEDLVMLPTPVRLTLNVLDDVLENEDEVEDVDEKCEQYGVVSL 360
DB 301 SPFFSIPFAPHIEDLVMLPTPVRLTLNVLDDVLENEDEVEDVDEKCEQYGVVSL 360
QY 361 LVPKEMRGQGVVEYANAGDSKAAQKLTGRMFDPKPVVATFYPLSAVKGRLYTTL 419
DB 361 LVPKEMRGQGVVEYANAGDSKAAQKLTGRMFDPKPVVATFYPLSAVKGRLYTTL 419

RESULT 8
ABP64662
ID ABP64662 standard; protein; 137 AA.

XX ABP64662;
AC
XX
XX 04-NOV-2002 (first entry)
DT
XX
XX
DE Human ORF1032.

XX Cytostatic; Cardiac; Anti-allergic; Immunosuppressive; Vascular;
XX Antiinflammatory; gene therapy; human; ORFX; atherosclerotic plaque;
XX human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
XX Cancer; cardiovascular disease; allergy; autoimmune disease;
XX wound healing; blood coagulation disorder; inflammatory disorder.

XX Homo sapiens.
OS
XX
XX US2002082206-A1.
PN
XX
XX 27-JUN-2002.

XX 30-MAY-2001; 2001US-00867550.
PF
XX
XX 30-MAY-2000; 2000US-0208427P.

XX (LEACH) LEACH M D.
PA (MEHR) MEHRABAN F.
PA (CONL) CONLEY P B.
PA (TOPP) TOPPER J N.
PA (LAWD) LAW D.

XX Leach MD, Mehraban F, Conley PB, Topper JN, Law D;
XX WPI; 2002-626554/67.
DR N-PSDB; ABQ99225.
XX
XX

PT New polypeptide designated ORFX are present in human atherosclerotic cells
PT and are useful to prevent and treat ORFX-associated disorders including
PT cancer, allergy, wound healing or autoimmune, cardiovascular or
PT inflammatory disease.

PS Claim 10; SEQ ID NO 2064; 78pp; English.

XX The present invention relates to novel human ORFX polypeptides and their
CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences
CC were discovered in human atherosclerotic cells, in particular in platelets
CC and human umbilical vein endothelial cells (HUVEC) and are expressed in
CC many other tissues as well. Atherosclerotic cells are cells which have the
CC potential to develop atherosclerotic plaques. The ORFX polypeptides and
CC nucleic acids are useful for treating or preventing a pathological
CC condition associated with an ORFX-associated disorder, e.g. cancer,
CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood
CC coagulation disorders or inflammatory disorders. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from the USPTO web site at
CC segdata.uspto.gov/sequence.html?DocID=2002082206

XX Sequence 137 AA:

Query Match 21.9%; Score 483.5; DB 5; Length 137;
Best Local Similarity 80.0%; Pred. No. 1e-43;
Matches 96; Conservative 1; Mismatches 12; Indels 11; Gaps 1;

QY 63 GAAASAAYEYGRKERALEQOGHRNIVTLVGVFTIHPSPVPSRCLLELDVSVSL 122
DB 29 GRGAALFSGGLAFAVA-----VTLVGVFTIHPSPVPSRCLLELDVSVSL 77

QY 123 LYSSHQCGSMWMIQHCARDVLEALAFHHEGYVHADLPKRNILWSAENECFKLIDFGISF 182
DB 78 LYSSHQCGSMWMIQHCARDVLEALAFHHEGYVHADLPKRNILWSAENECFKLIDFGISF 137

RESULT 9
AAG38502
ID AAG38502 standard; protein; 638 AA.

XX AC AAG38502;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 47508.
XX DE Protein identification; signal transduction pathway; metabolic pathway;
XX DE hybridization assay; genetic mapping; gene expression control; promoter;
XX DE termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131499P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.

PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145226P.
PR 26-JUL-1999; 99US-0145274P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147260P.
PR 05-AUG-1999; 99US-0147262P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0148341P.
PR 12-AUG-1999; 99US-0148342P.
PR 13-AUG-1999; 99US-0148656P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149930P.
PR 23-AUG-1999; 99US-0149930P.

PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155133P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161935P.
PR	29-OCT-1999;	99US-0162142P.

Query Match	11.0%	Score 243.5	DB 3	Length 638
Best Local Similarity	25.9%	Pred. No. 1,4e-16		
Matches	106	Conservative	56	Mismatches 146, Indels 101, Gaps 17

QY	11	EPFRFLA	FGRLMVOSRLGSGSSASVYRVRC	CGNSP	PGAL	KOLFDP	PGTGA	AAASAAE	70
Db	249	DPDFEFK	NGKLYOLGKIGSGSSSEVHKVI	---	SSDCTI	YALKKIKUKGRD	YATA	---	301
QY	71	YGFRKER	ALAELOLQHRNIVTLVGVFTI	HPNSV	RSCLLELD	VSYS	-----		120
Db	302	YGFQGEI	GYLKKLKKKTNIIQ	-----	IDYEVI	DKTLLEVL	NGTMSNDGR	KVEDGF	354
QY	121	---LLYS	-----	SHGCSM	-----	---			
Db	355	IYVLEGE	IDLAHL	SKRMREI	BEGSDRT	IDENMLR	IFYWQ	LLQVNTI	HERIYHSOLK
QY	161	PRNTLMS	ANECFKLID	FGLSKE	KNODV	KYIQTD	-----	GYRAPE	ALQNCIAQ
Db	415	PANFL	---LV	RGLK	IDFGIA	-KAINS	DDTNI	QRDSQ	VTLSYMSPEAF
		415							

```

QY      215 SDTECSAVDLMSGLITILEMFSGKKLKHTRYSGCWKSNSSAITD--HIFASKAVVAAI 272
Db      470 NTIKCGSPSDIWSGLGILIQWYVG-RTTPADYKTFW-AKFXVITDPNHEITYNQNSNP-- 525
QY      273 PAYHLRLDISMLHDDPSRRIPAEMAKCSPPFSIPFAPH-----311
Db      526 ---WLDLMKKCLAMDRNQRMRIPELLQHPLAPl-PhEQVKTIKLFSLIAESGSD 581
QY      312 -----IEDLVMPTVLRLLNLVLDLDYVGNSEEEYEDVDEVKEC 351
Db      582 DKANSMISQLEQLSNPAPLPr-----NDVLSDSQDQQLSKVSELC 624

```

RESULT 10	
AAg38501	
ID	AAg38501 standard; protein; 720 AA.
XX	
AC	AAg38501;
XX	
DT	18-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 47507.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

OS *Arabidopsis thaliana*

PN EP1033405-A2

PD 06-SEP-2000

PF 25-FEB-2000; 2000EP-00301439.

PR 25-FEB-1999; 99US-0121825P.

PR 09-MAR-1999; 99US-0123548P.

PR 25-MAR-1999; 99US-0126264P.

01-APR-1999; 99US-0127462P.

08-APR-1999; 99US-0128714P.

PR 19-APR-1999; 99US-0130077P.

23-APR-1999; 99US-0130510P.

28-APR-1999; 99US-0131449P.

30-APR-1999; 99US-0132407P.
PR

05-MAY-1999; 99US-0132485P.

06-MAY-1999: 99US-0132487P

PR 11-MAY-1999: 99US-0134256P

PR 14-MAY-1999: 99US-0134218P

PR	14-MAY-1999;	99US-013422LE
PR	14-MAY-1999;	99US-013437OP

PR	18-MAY-1999;	99US-0134768P.
DP	19-MAY-1999:	99IT-0134941P

PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.

PR 24-MAY-1999; 9905-0135629P.
 DD 25-MAY-1999; 9905-0136031P.

PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.

PR 01-JUN-1999; 99US-0137222P.

PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 14-JUN-1999; 99US-0138847P.
PR 16-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 18-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0140695P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139617P.
PR 21-JUN-1999; 99US-0139617P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-014187P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142290P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142820P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.

PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159332P.
PR 14-OCT-1999; 99US-0159333P.
PR 14-OCT-1999; 99US-0159334P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 11.0%; Score 243.5; DB 3; Length 720;
Best Local Similarity 25.9%; Pred. No. 1.7e-16;

PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145911P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148584P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149829P.
PR 20-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 05-OCT-1999; 99US-0157855P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 14-OCT-1999; 99US-0159644P.
PR 18-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.

PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161932P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 11.0%; Score 243.5; DB 3; Length 777;
Best Local Similarity 25.9%; Pred. No. 1.9e-16;
Matches 106; Conservative 56; Mismatches 146; Indels 101; Gaps 17;

QY 11 EPPRFLEAAGRIMQVQSRIGSGSSASVYRRCNPGSPGALKQFLPRTTGAASAAE 70
DB 388 DPDFFKVNGKLYQRIKIGSGSSSEVHKVI--SSDCTIYALKIKLGRDYATA--- 440
QY 71 YGFRKERAALEQLOQHRNIVTLVGYFTIHFSBNVPSRCILLELDVSYSE----- 120
DB 441 YGFCOEIGLYKTKLKKRTNIQL-----IDYEVTDKTLQELVANGTSMKQGRVKEGDF 493
QY 121 ---LLIYS---SHQCSNW-----MIQHCARDVLEALAFHEGYHADIK 160
DB 494 IYVLEVEYGEIDLAHMLSQKREIEGSDRTIDENMWRIFYWOQTLQAVNTIHERIYHSDIK 553
QY 161 PNIILMSANESCFKIDPELSFKENQDVKYQTD-----GYRPAELQWCLAQAGQ 214
DB 554 PANFL--LVRGFLKIDFGLA--KAINSDTNIQDSQVGTLSYMSPEAFMCNIESDENG-- 608
QY 215 SDTECTSAVDLWLSGLILEMFSGMKIKHTVRSQEMKANSSAIIID--HIFASKAVVNNAI 272
DB 609 NTKICGRPSDINSLGICILYQWYVG--RTPRADYKTFW--AKFKYITDPNHEITNQSLNP-- 664
QY 273 PAYHLRDLIKSMHDDPSRRIEAMALGSPFSPAPAH----- 311
DB 665 ---WLIDLMKICLAWDRNQWRIRIPELQHPFLAPPI--PHEPOVKTIKLFSLIAESCGSD 720
QY 312 -----IEDLVMLPTPVRLANLVDDDYGNEEVEDYEDYKEC 351
DB 721 DKANSMSIQLEQLLSNPAPLP-----NDVLDSRDQNOQLSRVSELC 763

RESULT 12
AAW85028
ID AAW85028 standard; protein; 544 AA.
XX
AC AAW85028;
XX
DT 17-OCT-2003 (revised)
DT 08-FEB-1999 (first entry)
XX
DE CDK2-green fluorescent protein fusion product.
XX
KW Human; CDK2 gene; fusion protein; green fluorescent protein; GFP;
KW intracellular signaling; chimera.
XX
OS Nequorea victoria.
OS Homo sapiens.
OS Chimeric.
XX
PN MO9845704-A2.
XX
PD 15-OCT-1998.
XX
PF 07-APR-1998; 98WO-DK000145.
XX
PF 07-APR-1997; 97DK-00000392.

XX (NOVO) NOVO-NORDISK AS.
 PA Thastrup O, Petersen Bjorn S, Tullin S, Kasper A, Scudder K;
 PI WPI; 1998-594491/50.
 DR N-PSDB; AAV71073.
 XX
 XX Determining effect on signalling pathways in live cells from
 PT redistribution of luminothores - specifically fusions of green
 PT fluorescent protein with a signalling component, and new apparatus,
 PT particularly for identifying toxins and potential therapeutic agents.
 XX
 XX Example 12; Page 208-209; 326pp; English.
 XX
 XX The present sequence represents a human CDK2-green fluorescent protein
 CC (GFP) fusion protein. The fusion protein is used in an assay to exemplify
 CC the invention. The specification describes how quantitative information
 CC about the influence of a molecule on a cellular response is obtained by
 CC recording the variation, caused by the molecule, on mechanically intact
 CC living cells, in the spatially distributed light emitted from a
 CC luminothore present in the cells. The variation in light emission is
 CC processed to provide information that correlates spatial distribution to
 CC the degree of the molecule. The method is used to identify agents that
 CC (in)directly affect intracellular signalling, especially to screen for
 CC potential therapeutic agents or toxins, and to identify new drug targets.
 CC (Updated on 17-Oct-2003 to standardise OS field)
 XX
 XX Sequence 544 AA;
 SQ
 Query Match 10.9%; Score 239.5; DB 2; Length 544;
 Best Local Similarity 27.0%; Pred. No. 3e-16;
 Matches 112; Conservative 52; Mismatches 154; Indels 97; Gaps 18;
 QY 23 MGVQSLGSGSSASVYRVRCGNGSPGALKQF-LPFGTTGAASAAEYGRKXRALE 81
 DB 4 FQKVEKIGEGTGVVYKAR--NKLTEGVVALKKIRLDTEBGPSTAI-----REISLTK 56
 QY 82 QLOGHNIYTLVGVFTIHSPVNSRCLLELDVSVSELLLYSSHOGCSMMIIOHCARD 141
 DB 57 EL-NHPIVLYKLDV--IHTENKL---YLVEFLHQDLKKRMDASAVTGIPPLIKSYLQ 110
 QY 142 VLEALAFLEHGEVYHADLKPRNIIWSAENECEFKLIDFGISFKEGNODVXY---IQDGYR 198
 DB 111 LLQGLAFCHSHRVLRHDLKPNLLINTEG-AIKLADFGIARAFGVPRVTYTHEVVTLMWR 169
 QY 199 APEAEIQLNCLAGLOSDTECTSAVDLSIGIILLFEMPSGMUKHTVRSQEKANSATI 258
 DB 170 APEI-----LLGSKYSTAVIDWSLGCIFAE-----VTRRALPPGDS--I 209
 QY 259 DHIFASKAVVN-----AIPAYHL-----RDLIKSLMD 287
 DB 210 DQLFRIFRLTGTDEVWVPGVTSMPDYKSPFKMARQDPSKVVPPLDEGRSLLSQMLHY 269
 QY 288 DPSRRIPAEMLCSPPFSIPFAPIEDLVMLPTPVRLNLVLDLDYLGNEEYEDVEDV 347
 DB 270 DEPKRISAKAALAHPPF-----QD-VTKVPBHLRLMD-----PVATWV 307
 QY 348 KEBCQCKGPRVSLVLPKENVGRGVFEVYANAGDSIAQKLITGRMFDGKFVAT 402
 DB 308 SKGEELFTGVVPLVELDGVNGHKF-SVSGEGEGDGYGKLT-----LKFICTT 356
 RESULT 13
 ADD21398 standard; protein; 298 AA.
 ID ADD21398 standard; protein; 298 AA.
 AC ADD21398;
 XX
 XX 15-JAN-2004 (first entry)
 DT
 XX Golden hamster cyclin dependent kinase (cdk) 2 protein.
 DE
 XX

KM continual growth; cultured cell; cyclin dependent kinase; cdk4; cdk2;
 KM cdk6; activating mutation; cell growth; cell division; cell cycle;
 KM cancer-causing agent; continual growth-induced cell; enzyme;
 KM golden hamster.
 XX
 XX Mesocricetus auratus.
 OS
 XX WO2003044169-A2.
 PN
 XX 30-MAY-2003.
 PD
 XX
 XX 15-NOV-2002; 2002MO-US036729.
 PF
 XX 15-NOV-2001; 2001US-0334760P.
 PR
 XX (UTEM) UNIV TEMPLE.
 PA Reddy PE, Rane SG, Mettus RV;
 XX WPI; 2003-449813/42.
 XX N-PSDB; ADD21399.
 DR
 XX
 XX A composition for reversibly inducing continual growth in normal cells
 PT comprises a cyclin dependent kinase protein (e.g. cdk4, cdk2 or cdk6) or
 PT its active fragment, derivative, homolog or analog, having an activating
 PT mutation.
 XX
 XX Disclosure; Page 104; 77pp; English.
 PS
 XX This invention relates to a novel composition for inducing a reversible
 CC state of a continual growth in cultured cells and comprises at least one
 CC compound comprising a cyclin dependent kinase (cdk4, cdk2 or cdk6
 CC protein having an activating mutation. Growth and division of living
 CC cells involve a regular series of events and processes that comprises the
 CC cell cycle. Cyclin dependent kinases cdk2, cdk4 and cdk6 are involved in
 CC the control of G1, the point at which cells irreversibly commit to DNA
 CC synthesis and thus enter the cell cycle. The invention is useful in
 CC reversibly inducing continual growth in normal cells and may allow the
 CC screening of cancer-causing agents with the continual growth-induced
 CC cells. The present sequence is that of the golden hamster cdk2 protein
 CC related to the invention. Note: Due to an error in the specification or
 CC sequence listing, the Seq ID numbers given in the disclosure do not
 CC correspond to those given in the sequence listing. It is therefore
 CC unclear which Seq ID number corresponds to which sequence and exactly
 CC which sequence is being claimed.
 XX
 XX Sequence 298 AA;
 SQ
 Query Match 10.8%; Score 238; DB 7; Length 298;
 Best Local Similarity 29.3%; Pred. No. 1.7e-16;
 Matches 99; Conservative 44; Mismatches 117; Indels 78; Gaps 16;
 QY 23 MGVQSLGSGSSASVYRVRCGNGSPGALKQF-LPFGTTGAASAAEYGRKXRALE 81
 DB 4 FQKVEKIGEGTGVVYKAR--NKLTEGVVALKKIRLDTEBGPSTAI-----REISLTK 56
 QY 82 QLOGHNIYTLVGVFTIHSPVNSRCLLELDVSVSELLLYSSHOGCSMMIIOHCARD 141
 DB 57 EL-NHPIVLYKLDV--IHTENKL---YLVEFLHQDLKKRMDASAVTGIPPLIKSYLQ 110
 QY 142 VLEALAFLEHGEVYHADLKPRNIIWSAENECEFKLIDFGISFKEGNODVXY---IQDGYR 198
 DB 111 LLQGLAFCHSHRVLRHDLKPNLLINTEG-AIKLADFGIARAFGVPRVTYTHEVVTLMWR 169
 QY 199 APEAEIQLNCLAGLOSDTECTSAVDLSIGIILLFEMPSGMUKHTVRSQEKANSATI 258
 DB 170 APEI-----LLGSKYSTAVIDWSLGCIFAE-----VTRRALPPGDS--I 209
 QY 259 DHIFASKAVVN-----AIPAYHL-----RDLIKSLMD 287
 DB 210 DQLFRIFRLTGTDEVWVPGVTSMPDYKSPFKMARQDPSKVVPPLDEGRSLLSQMLHY 269
 QY 288 DPSRRIPAEMLCSPPFSIPFAPIEDLVMLPTPVRLNLVLDLDYLGNEEYEDVEDV 347
 DB 270 DEPKRISAKAALAHPPF-----QD-VTKVPBHLRLMD-----PVATWV 307
 QY 348 KEBCQCKGPRVSLVLPKENVGRGVFEVYANAGDSIAQKLITGRMFDGKFVAT 402
 DB 308 SKGEELFTGVVPLVELDGVNGHKF-SVSGEGEGDGYGKLT-----LKFICTT 356
 QY 288 DPSRRIPAEMLCSPPFSIPFAPIEDLVMLPTPVRL 325

CC neuromuscular junction or sarcolemma needs to be disrupted
XX
SQ Sequence 367 AA;

Query Match 10.7%; Score 235.5; DB 3; Length 367;
Best Local Similarity 22.8%; Pred. No. 4.4e-16;
Matches 96; Conservative 67; Mismatches 127; Indels 131; Gaps 19;

```
QY 10 AEPRLLEAFGR-----LMQVQS-----RLGSSSASVYRRCGNGSPRG---ALK 54
DB 2 SSPPARKGPRQEVTKTAMEVRAVYQDLOPVSGSAGAV---CSAVDSRTGNKVAIK 56
QY 55 OFLPFGTGAASAAYGFRKERALELOLGHRIYTLYGVFTHFSFNVPSRC----- 108
DB 57 KLYRPFQSELFAKRA---YRELRLKHKR-HENVIGLDVFT-----PDETLDFTD 104
QY 109 --LLELLDVSVSELLLYSSHQCSMMMIQHCAVDYLEALAFHHEGYVHADIKPRNII 166
DB 105 FYLVMPFMGTDLGKLM---KHETLSEDRIOFLVYQMLKGIYHAGVTHRDLPGNL-- 159
QY 167 SAENEC-FKLIDRGLSPKGNQDVKIYIQDGYRAPAEALQNCIAQAGLSDTECTSAVDL 225
DB 160 AVNEDELKILDPGLARQADSEMTGYVTRKWRPAPEVIL-----NMKRYTQTVDI 209
QY 226 WSLGIIILEMFSGMKLKHTVRSQEMRKANSALIDHIFASKAVVNA----- 270
DB 210 WSYGCIAMEMITGKIL-----FKGN-----DHLQLEIKVKYGTGTPPEPFVQKLSA 256
QY 271 -----AIPAYHLRD-----LKSMLHDDPSRRIIPAEMLCSPPFSIPFA 309
DB 257 EAKNYMEGLPELEKDPASVLTNASPOAVNLEKMLVLDAEQRTAAEALAHPEFE--- 312
QY 310 PHIEDLVMLPTVLRLLNVLDDDYLGNEEYEDVVEDVK---EECOK--YGPVVSLLVPR 364
DB 313 -----SLRDTEDPRKQKIDSPDDVDRTLSEMKRVITYKEVLSFKPFR 355
QY 365 E 365
DB 356 Q 356
```

Search completed: November 29, 2004, 11:30:36
Job time : 112.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2004, 11:23:30 ; Search time 5310 Seconds

(without alignments)
8646.716 Million cell updates/sec

Title: US-10-798-532-3

Perfect score: 1260
Sequence: 1 atcgcgcgagctccgctgcgc.....tgatcaaaccttgcttaa 1260

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865308 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1113.4	88.4	1763	3	AK013347	AK013347 Mus muscu
2	886.6	70.4	2846	3	AK030152	AK030152 Mus muscu
3	861.8	68.4	1079	4	BM468107	BM468107 AGENCOURT
4	723.8	57.4	917	4	BM451184	BM451184 AGENCOURT
5	622.8	49.4	1015	6	BY112344	BY112344 BY112344
6	619.2	49.1	657	2	BS547675	BS547675 602575520
7	617.8	49.0	889	2	BS895119	BS895119 601436041
8	616.4	48.9	772	4	BI562420	BI562420 603256316
9	608.8	48.3	932	5	BU541306	BU541306 AGENCOURT
10	595	47.2	875	4	BI259568	BI259568 602968330
11	591.6	47.0	834	7	CF593363	CF593363 AGENCOURT
12	583.8	46.3	906	4	BS386905	BS386905 602454704
13	583.6	46.3	682	4	BI133581	BI133581 602997069
14	578	45.9	888	5	BQ225071	BQ225071 AGENCOURT
15	560	44.4	602	7	CO737403	CO737403 SLIT02C17
16	557.6	44.3	779	5	BU447831	BU447831 603767571
17	556.4	44.2	876	5	BU184346	BU184346 AGENCOURT
18	549.2	43.6	844	4	BS474616	BS474616 602517423
19	519.8	41.3	824	2	BS568528	BS568528 601342425
20	516.8	41.0	823	7	CO771206	CO771206 testis ES
21	516	41.0	781	7	CN232453	CN232453 WB074D11
22	498.6	39.6	717	5	BU388761	BU388761 603582650
23	497.8	39.5	670	7	CN359673	CN359673 170004245
24	496.8	39.4	523	2	BF929410	BF929410 IL2-NT020

25	495.2	39.3	896	7	CO773091	CO773091 testis ES
26	494.2	39.2	1003	5	BQ922020	BQ922020 AGENCOURT
27	493	39.1	928	5	BQ424438	BQ424438 AGENCOURT
28	482.4	38.3	716	7	CF745774	CF745774 UI-M-GVO-
29	471.8	37.4	755	5	BU731858	BU731858 603568233
30	469.8	37.3	794	5	BU205357	BU205357 603104559
31	468.2	37.2	824	4	BI253163	BI253163 602973284
32	467.2	37.1	785	6	CB990192	CB990192 AGENCOURT
33	463.6	36.8	870	5	BQ216813	BQ216813 AGENCOURT
34	463.4	36.8	484	5	BX281211	BX281211 UI-M-FRO-
35	463.4	36.8	699	5	BU707284	BU707284 UI-M-FRO-
36	460	36.5	696	5	BU439278	BU439278 604147528
37	456.6	36.2	483	2	BF299411	BF299411 IL2-NT020
38	454.2	36.0	631	5	BU112717	BU112717 603131421
39	451.8	35.9	743	4	BI562440	BI562440 603256324
40	450.4	35.7	986	5	BU751496	BU751496 CH4#002.E
41	449	35.6	933	2	BS886040	BS886040 AGENCOURT
42	433.2	34.4	820	7	CK597557	CK597557 AGENCOURT
43	427.4	33.9	672	5	BX261106	BX261106 BX261106
44	420	33.3	604	5	BX261107	BX261107 BX261107
45	413.2	32.8	570	5	BU293329	BU293329 60416812

ALIGNMENTS

RESULT 1
AK013347
LOCUS 1763 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length
enriched library, clone:281045404 product:kinase interacting with
leukemia-associated gene (statmin), full insert sequence.

ACCESSION AK013347 GI:12850651
VERSION AK013347
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclirognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Komano, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 11042159
PUBMED 20499374

REFERENCE
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Komano, H., Akiyama, J., Nishi, K., Katsunai, T., Tachino, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujisawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M.,
Okazaki, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Yoneda, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

TITLE RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multichannel sequencer
Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861

REFERENCE
AUTHORS THE RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

REFERENCE
AUTHORS THE FANTOM Consortium and the RIKEN Genome Exploration Research

Db	1239	CTGTGTCAAAGAAAAATCCCTGGCAGACGACAAGTCTTGTAAGTACGCACACGCTGCT	1298
Oy	1141	GATTCCAAAGCTGGCGCAGAATACTGACTGGAAGATGTTTGATGGGAAGTTGTTGTG	1200
Db	1299	GATTCCAAAGCTGCTCAAGAACTGCTGACTGGGAGAGATGTTGACGGGAAATTTGTTGTG	1358
Oy	1201	GCTACATCTCTACCCGCTGAGGCCCTACAGAGGGAGATTCGTATTCAAACCTTGCTTAA	1266
Db	1359	GCTACATCTCTACCCGCTGAGGCCCTACAGAGGGAGATTCGTATTCAAACCTTGCTTAA	1418
RESULT 2			
AKO30152			
LOCUS	AKO30152	2846 bp mRNA linear HTC 03-APR-2004	
DEFINITION	Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:493401D07 product:kinaase interacting with leukemia-associated gene (stathmin), full insert sequence.		
ACCESSION	AKO30152		
VERSION	AKO30152.1 GI:26326136		
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999) 99279253		
JOURNAL MEDLINE PUBMED	103496636		
AUTHORS	2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)		
JOURNAL MEDLINE PUBMED	20499374		
REFERENCE	11042159		
AUTHORS	3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Kono,H., Akiyama,J., Nishi,K., Kitenuai,T., Tashiro,H., Itoh,M., Sumi.N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsunoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujikake,S., Inoue,K., Togawa,Y., Iwama,M., Ohara,E., Matshiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawat,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kiru,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)		
JOURNAL MEDLINE PUBMED	20530913		
REFERENCE	11076861		
AUTHORS	4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)		
TITLE	5		
JOURNAL REFERENCE	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002) 6 (bases 1 to 2846)		
AUTHORS	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hizumoto,K., Hirooka,T., Hiroyane,T., Horii,F., Imocani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Kono,H., Kouda,M., Koya,S., Kuribara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi.K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saio,C., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shihagawa,A., Shiraki,T.,		

TITLE	Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.
JOURNAL	Direct Submission Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/. Location/Qualifiers
FEATURES	Source
	1..2846 organism="Mus musculus" mol_type="mRNA" strain="C57BL/6J" db_xref="FANTOM DB:4933401D07" db_xref="taxon:10090" clone="4933401D07" sex="male" tissue_type="testis" cdon_lib="RIKEN full-length enriched mouse cDNA library" idev_stage="adult" size..1230 note="unnamed protein product; kinase interacting with leukemia-associated gene (stathmin) [MGI:MGI:1341908] putative"
CDS	/codon_start=1 protein_id="BAC26812.1" db_xref="GI:26326137" translation="MTLYGVFTIHFSBNVSRCLLELDVSVELLYSHHOCSSMWI MICHRADYVEALAFTHIEGYVHADLPRIILMSABNECPKIDFGISFKEGQDVXY IQDVGVAPEAEONCLAGOSDFCTSAVDLMSIGILLMEPSMKRKHVRSOE WKNSSAIIDHIFASRAVVNAAPAYHAILITSMHDDGRRIIPAMALCSPEFI PAHHIEDLVMLFPVRLNLAVLDDDIEMEDRIDVEDYKRCQKRGVSVLSVPKE NPERGVFVRYANAGSKAAOKLLTGRMPGKFVAVTFPLSAVKRGIVYTLL" 2833..2828 note="putative" 2846 note="putative"
ORIGIN	polya_signal polya_site note="putative"
	Query Match 70.4%; Score 886.6; DB 3; Length 2846; Best local Similarity 93.1%; Pred. No. 8.8e-247; Matches 928; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
D1	264 CATCGATGACTTTGTATGGAGTGTTTACATTCACCTTTCTCCAAATGTGCATTAGCGTG 323
D2	234 CTTCATGACTTTAATGATGAGTCTTTTACCATACTTCTCTCCCAATGTGCATCACCGTG 293
D3	324 TCCTGGCTGGAACCTCGATATGCACTGTTTGGAAATTCGCTTAATATTCACAATCCA 383
D4	294 TCTGCTCTGTAACCTCTGGAATGCACTGTTTGGAAATTCGCTTAATATTCACAATCCA 353
D5	384 GGGTGTTCATGATGATGATACAGACTTCGCCGCCGAGATGTTTTGGAGGCGCTTGCTTT 443
D6	354 GGGGTGCTCATATGATGATGATACAGACTTCGCCGCCGAGATGTTTTGGAGGCGCTTGCTTT 413
D7	444 TCTTCATCATGAGGGCTATGTCCATGCGGACCTCAACACGTAACATATTTGGAATGC 503
D8	414 TCTTCACCAAGAAGGCTATGCTCATGACAGACCTCAACACGAAACATCTCTGGAATGC 473
D9	504 AGAAGAAAGATGTTTTAACTCATGACTTTGGACTTAAGCTTCAAAGAAGCAATCAGGA 563
D10	474 GGAGAAAGAGGTTTTAAGCTTATATGACTTTGACTCAGCTTCAAAGAAGCAATCAGGA 533

QY 1047 GGAGTGTCAAAATATGACCAAGTGTATCTTACTTGTTCAGAAATCTGGCAG 1106
 |||||
 DB 781 GGAGTGTCAAAATATGACCAAGTGTATCTTACTTGTTCAGAAATCTGGCAG 840
 |||||
 QY 1107 AGGACAAGTC-TTGTGTAGTATGCAATCTG-GTATTCGAAAGCTGGCAAAAT 1163
 |||||
 DB 841 AGGACAAGTC-TTGTGTAGTATGCAATCTG-GTATTCGAAAGCTGGCAAAAT 899
 |||||

RESULT 4
 BM451184 917 bp mRNA linear EST 05-FEB-2002
 LOCUS AGENCOURT 6392719 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5495159
 DEFINITION 5', mRNA sequence.
 ACCESSION BM451184
 VERSION BM451184.1 GI:18500224
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 917)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgsdbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM12121 row: j column: 24
 High quality sequence stop: 718.
 Location/Qualifiers
 1. 917
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5495159"
 /issue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: PCW-SPORT6; Site: 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.75 kb. Library constructed by Life
 Technologies."

ORIGIN

Query Match 57.4%; Score 723.8; DB 4; Length 917;
 Best Local Similarity 99.7%; Pred. No. 26-199;
 Matches 725; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 534 TGCATCTAGCTTAAAGAGGCAATCAGATGTAAAGTATATTCAGACAGCGGATCG 553
 |||||
 DB 21 TGGGCTTAGCTTAAAGAGGCAATCAGATGTAAAGTATATTCAGACAGCGGATCG 80
 |||||

QY 594 GGCTTCAGAAAGCAATGTGCAAAATGCTGGCCAGGCTGGCTGAGAGTATCAGA 653
 |||||
 DB 81 GGCTTCAGAAAGCAATGTGCAAAATGCTGGCCAGGCTGGCTGAGAGTATCAGA 140
 |||||

QY 654 ATGTACTCAGCTGTGTATCTGTGAGCCTAGGATCATTTTACTGGAAATGTTCTCAG 713
 |||||
 DB 141 ATGTACTCAGCTGTGTATCTGTGAGCCTAGGATCATTTTACTGGAAATGTTCTCAG 200
 |||||

QY 714 AATGAACTAAACATACATCTGATCTGAGAAATGGAAGCAACAGTTCTCTATAT 773
 |||||
 DB 201 AATGAACTAAACATACATCTGATCTGAGAAATGGAAGCAACAGTTCTCTATAT 260
 |||||

QY 774 TGATCACAATTTTGCAGTAAGAGAGTGTGAATGCCGAATTCAGACTATACCTAAG 833
 |||||

DB 261 TGATCACAATTTTGCAGTAAGAGAGTGTGAATGCCGAATTCAGACTATACCTAAG 320
 |||||
 QY 834 AATCCTTAAAGAGAGTGTGAATGCCGAATTCAGACTATACCTAAG 893
 |||||
 DB 321 AATCCTTAAAGAGAGTGTGAATGCCGAATTCAGACTATACCTAAG 380
 |||||

QY 894 ATTGTACAGCCCATCTTTAGCATTCCTTTTGGCCCTCATATGGAATCTGGTCACT 953
 |||||
 DB 381 ATTGTACAGCCCATCTTTAGCATTCCTTTTGGCCCTCATATGGAATCTGGTCACT 440
 |||||

QY 954 TCCCACTCAGTGTCTAAGAGCTGTGAATGTGTGATGATGATTTCTTGGAAATGAGA 1013
 |||||
 DB 441 TCCCACTCAGTGTCTAAGAGCTGTGAATGTGTGATGATGATTTCTTGGAAATGAGA 500
 |||||

QY 1014 GGAATATGAAAGTGTGTGAAAGTGTAAAGAGAGTGTCAAAATATGACCAAGTGT 1073
 |||||
 DB 501 GGAATATGAAAGTGTGTGAAAGTGTAAAGAGAGTGTCAAAATATGACCAAGTGT 560
 |||||

QY 1074 ATCTCTACTGTGTCCAAAGAGAAATCTGGCAGAGACAAAGTCTTGTGAGTATGACAA 1133
 |||||
 DB 561 ATCTCTACTGTGTCCAAAGAGAAATCTGGCAGAGACAAAGTCTTGTGAGTATGACAA 620
 |||||

QY 1134 TGTGTGTATTCGCAAGTGTGCGCAGAAATTAATCTGACCTGAGAGATGTTGAGGAA 1193
 |||||
 DB 621 TGTGTGTATTCGCAAGTGTGCGCAGAAATTAATCTGACCTGAGAGATGTTGAGGAA 680
 |||||

QY 1194 TGTGTGTATTCGCAAGTGTGCGCAGAAATTAATCTGACCTGAGAGATGTTGAGGAA 1253
 |||||
 DB 681 TGTGTGTATTCGCAAGTGTGCGCAGAAATTAATCTGACCTGAGAGATGTTGAGGAA 740
 |||||

QY 1254 GCTTTAA 1260
 |||||
 DB 741 GCTTTAA 747
 |||||

RESULT 5
 BY712344 1015 bp mRNA linear EST 17-DEC-2002
 LOCUS BY712344
 DEFINITION Mus musculus cDNA clone 2810454004 5', mRNA sequence.
 ACCESSION BY712344
 VERSION BY712344.1 GI:27123621
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 REFERENCE 1 (bases 1 to 1015)
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamataka, I.,
 Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
 Schonbach, C., Gotohori, T., Baldairelli, R., Hill, D.P., Bulic, C.,
 Hume, D.A., Quackenbush, J., Schrim, L.M., Kanapin, A., Matsuda, H.,
 Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,
 Chochia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
 Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
 Gariboldi, M., Gisel, C., Godzik, A., Gough, J., Grimmond, S.,
 Gustincich, S., Hirokawa, N., Jackson, J., Jarvis, B.L., Konagaya, A.,
 Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
 Kurukin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
 Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
 Numata, K., Okido, T., Pavan, W.U., Pereira, G., Pesole, G.,
 Petrovsky, N., Pillai, R., Pontius, J.U., Q.D., Ramachandran, S.,
 Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
 Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shinada, K.,
 Sultana, R., Takemura, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
 Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
 Wells, C., Wilming, L.G., Wymshaw-Borja, A., Yanagisawa, M., Yang, I.,
 Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
 Hayatsu, N., Hirozane-Kobayashi, T., Kono, H., Nakamura, M.,
 Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,

TITLE
JOURNAL
MEDLINE
PUBMED

COMMENT

Itch,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
Shingawa,A., Yasuniishi,A., Yoshino,M., Waterston,R., Lander,B.S.,
Rogers,J., Birney,E. and Hayashizaki,Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

22354683
12466851
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@sc.riken.jp, URL:http://genome-gsc.riken.jp/
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,
Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F.,
Imotani,K., Ishii,Y., Itch,M., Kagawa,I., Kawai,J., Kojima,Y.,
Kondo,S., Kono,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,
Nomura,K., Numazaki,R., Ohno,M., Ohnato,N., Saito,R., Sakazume,N.,
Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M.,
Takeda,Y., Waki,K., Wataniki,A., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedic: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (http://genome-gsc.riken.go.jp) for
further details.

FEATURES

source

Location/Qualifiers
1..1015
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="2810454004"
/tissue_type="whole body"
/dev_stage="10, 11 days embryo"
/clone_lib="RIKEN full-length enriched, 10, 11 days embryo
whole body"

ORIGIN

Query Match 49.4%; Score 622.8; DB 6; Length 1015;
Best Local Similarity 90.0%; Pred. No. 6,3e-170;
Matches 678; Conservative 0; Mismatches 73; Indels 2; Gaps 1;

QY 1 ATGGCGGGAATCCGGCTGGCGTGGGCGGAGACCCCGCTTTTCTGGAGGCGCTTCGGG 60
DB 159 ATGGCGGGAATCCGGCTGGCGTGGGCGGAGACCCCGCTTTTCTGGAGGCGCTTCGGG 218
QY 61 CGGCTGTGGCAGGTACAGAGCCGTGTGGTGGGCGGCTTCGGCTCGGATGTGGGGTT 120
DB 219 CGGCTGTGGCAGGTACAGAGCCCGCTGGGCGAGCGGCTTCGGCTTCGGATACCGGGTG 278
QY 121 CGCTGTGGCGCAACCTGTGGCTGGCCCGCGGCGCCCTCAGGAGATTCTTGC CGCAGGA 180
DB 279 CGGCTGTGGCGGTACCGCGGCGCTGGCCCGCGGCGCCCTCAGGAGATTCTTGC CGCAGGA 338
QY 181 ACCACCGGGGCTGGCGCTTCGGCGGAGTATGTTTCCGCAAGAAGAGGGGGCGCTG 240
|||||

DB 339 ACCACCGGGGCTGGCGCTTCGGCGGAGTATGTTTCCGCAAGAAGAGGGGGCGCTG 398
QY 241 GAACAGTTGCAAGGCTCAGCAAGAAATCGTATGTTATGAGTGTATCAATCACTTT 300
DB 399 GANCAAGTTGCAAGGCTCAGCAAGAAATCGTATGTTATGAGTGTATCAATCACTTT 458
QY 301 TCTCCAAATGTCGATCAGCGCTGTCTGTGCTTGAACCTCTGGATGTCAAGTTCGGAA 360
DB 459 TCTCCAAATGTCGATCAGCGCTGTCTGTGCTTGAACCTCTGGATGTCAAGTTCGGAA 518
QY 361 TTGCTCTTATATTCAGTACAGCAGGAGTGTTCAGTGGATGATCAGATTCGCGCCGA 420
DB 519 TTGCTCTTATATTCAGTACAGCAGGAGTGTTCAGTGGATGATCAGATTCGCGCCGA 578
QY 421 GATGTTTGGAGCGCTTGTCTTTCTTCATCATGAGGCTATGTCATGCGGACTCAAA 480
DB 579 GATGTTTGGAGCGCTTGTCTTTCTTCATCATGAGGCTATGTCATGCGGACTCAAA 638
QY 481 CCACTTACATATTTGTGAGTGCAGAGAAATGATTTTAACTCATTTGACTTTGACTT 540
DB 639 CCACGAAACATCTGTGAGTGCAGAGAAATGATTTTAACTCATTTGACTTTGACTT 698
QY 541 AGCTTCAAGAGGGAATCAGAGTGAATGAATATATTCAGACAGAGGGATGCGGCTCCA 600
DB 699 AGCTTCAAGAGGGAATCAGAGTGAATGAATATATTCAGACAGAGGGATGCGGCTCTT 758
QY 601 GAAGCAGATTTGCAAAATTTGCTTGGCCGAGCGCTGCTGCAAGTATCAATGATACC 660
DB 759 GAAGCAGATTTGCAAAATTTGCTTGGCCGAGCGCTGCTGCAAGTATCAATGATACC 818
QY 661 TGAAGCTTTGATCTGTGAGACCTAGAGAAATCATTTTACTGGAATTTCTCAGGAATGAA 720
DB 819 TGAAGCTTTGATCTGTGAGACCTAGAGAAATCATTTTACTGGAATTTCTCAGGAATGAA 878
QY 721 CTGAACATACAGTCAGATCTCAGGAATGGAAG 753
DB 879 CTGAACATACAGTCAGATCTCAGGAATGGAAG 909
|||||

RESULT 6
BG547675 657 bp mRNA linear EST 04-APR-2001
LOCUS
DEFINITION
602575520P1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4703645 5',
mRNA sequence.

ACCESSION
BG547675
VERSION
BG547675.1 GI:13546340
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-rcmail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNLI at:
http://image.llnl.gov
Plate: L1CM1542 row: g column: 06
High quality sequence stop: 657.

FEATURES

source

Location/Qualifiers
1..657
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4703645"
/lab_host="DH10B (TI phage-resistant)"

/clone.lib="NIH_MGC_77"
 /note="Organ: lung; Vector: PDNR-LIB (Clontech); Site 1:
 5'flr (ggcgccctggcc); Site 2: 5'flr (ggcattatggcc); 5' and
 3' adaptors were used in cloning as follows: 5' adaptor
 sequence: 5'-CAGGCCATTATGGCC-3' and 3' adaptor sequence:
 5'-ATTCTAGAGCCGCGCGCCGCGCAGT-dt(30)NH-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size 1.9
 kb (range 0.5-4.0 kb). 12/15 clones contained inserts
 by PCR. This library was enriched for full-length clones
 and was constructed by Clontech Laboratories (Palo Alto,
 CA). Note: this is a NIH_MGC library."

ORIGIN

Query Match 49.1%; Score 619.2; DB 4; Length 657;
 Best Local Similarity 99.1%; Pred. No. 6.2e-169;
 Matches 654; Conservative 0; Mismatches 3; Indels 3; Gaps 3;

QY 577 CAGACAGACGGGTATCGGGCTCCAGAGCAGAAATTCGTTGGCCAGGCTGGC 636
 DB 1 CAGACAGACGGGTATCGGGCTCCAGAGCAGAAATTCGTTGGCCAGGCTGGC 60
 QY 637 CTGCAGAGTATCAGAAATGTACTCAGCTGTGATCTGGAGCCTAGGAATCATTTTA 696
 DB 61 CTGCAGAGTATCAGAAATGTACTCAGCTGTGATCTGGAGCCTAGGAATCATTTTA 120
 QY 697 CTGGAATATGTTCTCAGAAATGAAACTGAAACATACAGTCACTCAGGAATGGAAGGA 756
 DB 121 CTGGAATATGTTCTCAGAAATGAAACTGAAACATACAGTCACTCAGGAATGGAAGGA 180
 QY 757 AACAGTTCTGCTATTATTCATCATATTTGCACTGTAAGCAGTGGTGAATGCCGCAAT 816
 DB 181 AACAGTTCTGCTATTATTCATCATATTTGCACTGTAAGCAGTGGTGAATGCCGCAAT 240
 QY 817 CCAAGCTTATCCTTAAGAGACTTATCAAAAGATGCTTCATGATGATCCAGCAGAGA 876
 DB 241 CCAAGCTTATCCTTAAGAGACTTATCAAAAGATGCTTCATGATGATCCAGCAGAGA 300
 QY 877 ATTCCTGCTGAATGAGCATTTGGCAGCCCATCTTTAGCATTCCTTTGCCCTCATATT 936
 DB 301 ATTCCTGCTGAATGAGCATTTGGCAGCCCATCTTTAGCATTCCTTTGCCCTCATATT 360
 QY 937 GAAGATCTGTCATGCTCCACTCCAGTCTAAGACTGCTGAATGCTGATGATGAT 996
 DB 361 GAAGATCTGTCATGCTCCACTCCAGTCTAAGACTGCTGAATGCTGATGATGAT 420
 QY 997 TATCTTGGAGTGAAGAGAAATGAAATGTTGTAAGATGTAAGAAGAGAGTCAA 1056
 DB 421 TATCTTGGAGTGAAGAGAAATGAAATGTTGTAAGATGTAAGAAGAGAGTCAA 480
 QY 1057 AAATATGACCAAGTATCTCTACTTGTCCAAAGAAATCTCTGGCAGAGCAAGTC 1116
 DB 481 AAATATGACCAAGTATCTCTACTTGTCCAAAGAAATCTCTGGCAGAGCAAGTC 540
 QY 1117 TTTGTTGATGATCAAAATCTGCTGATTCCAAAGCTGCGCAGAAATTAATCTGA 1176
 DB 541 TTTG-TGATGATGCAAAATCTGCTGATTCCAAAGCTGCGCAGAAATTAATCTGA 599
 QY 1177 ATGTTGATGGAAGTTGTTGCTGCTACATTTACCCGCTGAGTCCCTTACAAAGAGGA 1236
 DB 600 ATGTTGATGGAAGTTG-TGCTGCTACATTTACCCGCTGAGT-CTACAAAGAGGGA 657

RESULT 7
 BE895119 899 bp mRNA linear EST 20-OCT-2000
 LOCUS 60136041F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3920939 5',
 DEFINITION mRNA sequence.
 ACCESSION BE895119
 VERSION BE895119.1 GI:10358193
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
 1 (bases 1 to 899)
 NIH-MGC <http://mgc.ncbi.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgaabs-remail.nih.gov
 Tissue Procurement: ATCC/DCPD/DPF
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 plate: LLM9753 row: b column: 12
 High quality sequence stop: 626.
 Location/Qualifiers

FEATURES
 source
 1..899
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3920939"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone.lib="NIH_MGC_72"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 2 kb. Library constructed by Life
 Technologies."

ORIGIN

Query Match 49.0%; Score 617.8; DB 2; Length 899;
 Best Local Similarity 99.7%; Pred. No. 1.8e-168;
 Matches 619; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 337 CTCTGGATGTCAGTGTTCGGAATTCCTTAATTCAGTCAACAGGTTGTCATG 396
 DB 1 CTCTGGATGTCAGTGTTCGGAATTCCTTAATTCAGTCAACAGGTTGTCATG 60
 QY 397 TGGATGATACAGCATTCGCGCCGAGATGTTTGAAGCCCTTGTCTTTCATCATGAG 456
 DB 61 TGGATGATACAGCATTCGCGCCGAGATGTTTGAAGCCCTTGTCTTTCATCATGAG 120
 QY 457 GGGTATGTCATGCGGCACTCAACACGTAACTATTTGGAGTGCAGAGAAATGAAT 516
 DB 121 GGGTATGTCATGCGGCACTCAACACGTAACTATTTGGAGTGCAGAGAAATGAAT 180
 QY 517 TTTAACTCATTTGACTTTGGACTTAGCTTCAAGAGCAATCAGATGTAAGTATAT 576
 DB 181 TTTAACTCATTTGACTTTGGACTTAGCTTCAAGAGCAATCAGATGTAAGTATAT 240
 QY 577 CAGACAGACGGGTATCGGGCTCCAGAGCAGAAATTCGTTGGCCAGGCTGGC 636
 DB 241 CAGACAGACGGGTATCGGGCTCCAGAGCAGAAATTCGTTGGCCAGGCTGGC 300
 QY 637 CTGCAGAGTATCAGAAATGTACTCAGCTGTGATCTGGAGCCTAGGAATCATTTTA 696
 DB 301 CTGCAGAGTATCAGAAATGTACTCAGCTGTGATCTGGAGCCTAGGAATCATTTTA 360
 QY 697 CTGGAATATGTTCTCAGAAATGAAACTGAAACATACAGTCACTCAGGAATGGAAGGA 756
 DB 361 CTGGAATATGTTCTCAGAAATGAAACTGAAACATACAGTCACTCAGGAATGGAAGGA 420
 QY 757 AACAGTTCTGCTATTATTCATCATATTTGCACTGTAAGCAGTGGTGAATGCCGCAAT 816
 DB 421 AACAGTTCTGCTATTATTCATCATATTTGCACTGTAAGCAGTGGTGAATGCCGCAAT 480
 QY 817 CCAAGCTTATCCTTAAGAGACTTATCAAAAGCATGCTTCATGATGATCCAGCAGAGA 876
 DB 481 CCAAGCTTATCCTTAAGAGACTTATCAAAAGCATGCTTCATGATGATCCAGCAGAGA 540
 QY 877 ATTCCTGCTGAATGAGCATTTGGCAGCCCATCTTTAGCATTCCTTTGCCCTCATATT 936

```

Db      541 ATTCCTGCTGAATAGGATGTGACAGCCATCTTTAGATTCCTTTTGGCCCTCATATT 600
Qy      937 GAAGATCTGATGATGCTTCCC 957
Db      601 GAAGATCTGATGATGCTTCC 621

RESULT 8
B1562420
LOCUS    B1562420          772 bp    mRNA    linear    EST 05-SEP-2001
DEFINITION 60356316F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5298674 5',
            mRNA sequence.
ACCESSION B1562420
VERSION   B1562420.1 GI:15449747
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE NIH-MGC http://mgs.nci.nih.gov/.
AUTHORS   1 (bases 1 to 772)
           National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
           Email: cgabbs-r@mail.nih.gov
           Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
           CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
           Toshitaki and Piero Carninci (RIKEN)
           CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Incyte Genomics, Inc.
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           plate: LLAM11755 row: p column: 03
           High quality sequence stop: 727.
           Location/Qualifiers
             1..772
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:5298674"
               /lab_host="DH10B"
               /clone_1ib="NIH_MGC_97"
               /note="Organ: testis; Vector: pBluescript (modified
               pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
               (GTCGAG); Oligo-dT primed using primer
               5'-TTTTTTTTTTTTTTVN-3', size-selected for average
               insert size 2.2 kb and normalized to 10^5. This is a
               primary library enriched for full-length clones and
               constructed using the Cap-trapper method (Carninci, in
               preparation). Library constructed by M. Brownstein
               (NHGRI/NHGRI, National Institutes of Health). Note: this is
               a NIH_MGC Library."

ORIGIN
Query Match      48.9%; Score 616.4; DB 4; Length 772;
Best Local Similarity 99.4%; Pred. No. 4.3e-168;
Matches 628; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy      274 TTGTATGAGATGTTTACATCACTTTCTTCGAAATGTGCATCAGGCTGTCTTGTCTT 333
Db      108 TTGTATGAGATGTTTACATCACTTTCTTCGAAATGTGCATCAGGCTGTCTTGTCTT 167
Qy      334 GAACCTCGATGATGCTGTTGGAAATGCTCTTATATTCAGTACACGAGGTGTTC 393
Db      168 GAACCTCGATGATGCTGTTGGAAATGCTCTTATATTCAGTACACAGGTGTTC 227
Qy      394 ATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 453
Db      228 ATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 287
Qy      454 GAGGGCTATGTCATGCGGACCTTAAACAGTAACTATTTGAGTGCAGAGATGAA 513

```

```

Db      288 GAGGCTATGTCATGCGGACCTTAAACAGTAACTATTTGAGTGCAGAGATGAA 347
Qy      514 TGTTTTAAATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 573
Db      348 TGTTTTAAATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 407
Qy      574 ATTCAGACAGACGGGTATCGGGCTCCAGAGCAGAAATTTGCTTGGCCAGGCT 633
Db      408 ATTCAGACAGACGGGTATCGGGCTCCAGAGCAGAAATTTGCTTGGCCAGGCT 467
Qy      634 GGCCTGCGAGAGTATACAGAAATGTAACCTGAGCTGTGATCTGTGAGGCTAGAAATCAT 693
Db      468 GGCCTGCGAGAGTATACAGAAATGTAACCTGAGCTGTGATCTGTGAGGCTAGAAATCAT 527
Qy      694 TTAGTGAATATGTTCTCAGAGATGAAATGAAACATAGATGATGATGATGATGATGATGAT 753
Db      528 TTAGTGAATATGTTCTCAGAGATGAAATGAAACATAGATGATGATGATGATGATGATGAT 587
Qy      754 GCAACAGTTCGCTATTTATGATGACA-TATTTGCCAGTAAAGCAGTGTGATGCCGC 812
Db      588 GCAACAGTTCGCTATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 647
Qy      813 AATTCAGGCTATCACTTACAGAGACCTTATCAAAAGCATGCTTCATGATGATCAAGCAG 872
Db      648 AATTCAGGCTATCACTTACAGAGACCTTATCAAAAGCATGCTTCATGATGATCAAGCAG 707
Qy      873 AAGAAATCCTGCTGAATGCGATGTCAGCC 904
Db      708 AAGAAATCCTGCTGAATGCGATGTCAGCC 719

```

```

RESULT 9
B0541306          932 bp    mRNA    linear    EST 13-SEP-2002
LOCUS    B0541306
DEFINITION AGENCOURT_10327419 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6572970
            5', mRNA sequence.
ACCESSION B0541306
VERSION   B0541306.1 GI:22851747
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE NIH-MGC http://mgs.nci.nih.gov/.
AUTHORS   1 (bases 1 to 932)
           National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
           Email: cgabbs-r@mail.nih.gov
           Tissue Procurement: DCTD/DRP
           CDNA Library Preparation: Rubin Laboratory
           CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Agencourt Bioscience Corporation
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           plate: LICM2765 row: o column: 18
           High quality sequence stop: 631.
           Location/Qualifiers
             1..932
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:6572970"
               /issue_type="carcinoma, cell line"
               /lab_host="DH10B (phage-resistant)"
               /clone_1ib="NIH_MGC_40"
               /note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
               Site 2: EcoRI; cDNA made by oligo-dT priming.
               Directionally cloned into EcoRI/XhoI sites using the
               following 5' adaptor: GGCACAG(G). Library constructed by
               Ling Hong in the laboratory of Gerald M. Rubin (University
               of California, Berkeley) using ZAP-cDNA synthesis kit

```

FEATURES

source

(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

Query Match 48.3%; Score 608.8; DB 5; Length 932;
Best Local Similarity 93.7%; Pred. No. 7.5e-166;
Matches 689; Conservative 0; Mismatches 38; Indels 8; Gaps 5;

1 ATGGCGGAGTCCGGGCTGGGCGGCGGAGCCCGCGGTTTCTGAGGCGCTTCGGG 60
149 ATGGCGGAGTCCGGGCTGGGCGGCGGAGCCCGCGGTTTCTGAGGCGCTTCGGG 208
61 CGGCTGTGCGAGTACAGAGCCGCTGTGGTAGCGGCTCCGCTCGGTATTCGGGTT 120
209 CGGCTGTGCGAGTACAGAGCCGCTGTGGTAGCGGCTCCGCTCGGTATTCGGGTT 268
121 CGCTGTGCGGCAACCTTGCTGCCCCCGGCGCTTCAGGAGTTCTTGCGCCAGGA 180
269 CGCTGTGCGGCAACCTTGCTGCCCCCGGCGCTTCAGGAGTTCTTGCGCCAGGA 328
181 ACCACCGGGGCTGCGGCTCTGCGCGCGAGTATGTTCCGCAAGAGAGGGCGGCTG 240
329 ACCACCGGGGCTGCGGCTCTGCGCGAGTATGTTCCGCAAGAGAGGGCGGCTG 388
241 GAACAGTTCAGGAGTACAGAAACATCTGATCTTGTATGAGTGTTCACATTCACATT 300
389 GAACAGTTCAGGAGTACAGAAACATCTGATCTTGTATGAGTGTTCACATTCACATT 448
301 TCTCCAAATGTGCGCATCAGCTGTCTGTGTTGAACTCCGTGAGTCAAGTGTTCGAA 360
449 TCTCCAAATGTGCGCATCAGCTGTCTGTGTTGAACTCCGTGAGTCAAGTGTTCGAA 508
361 TTGCTCTTATTTCCAGTACACAGGTTGTTCCATGATGATATAGATTTGGCGGCA 420
509 TTGCTCTTATTTCCAGTACACAGGTTGTTCCATGATGATATAGATTTGGCGGCA 568
421 GATGTTTGGAGGCCCTTGTCTTTCATCATGAGGCTATGTCAGCGGACT-CAA 479
569 GATGTTTGGAGGCCCTTGTCTTTCATCATGAGGCTATGTCAGCGGACTCCCAA 628
480 ACCACGTAACATATGTGAGTGCAGAGATGAAATGTTTAACTCATTTGACTTGGAC- 538
629 ACCACGTAACATATGTGAGTGCAGAGATGAAATGTTTAACTCATTTGACTTGGAC 688
539 TTAGCTTCAAAGAGGCAATCAGATGTAAGTATTCAGACAGCGGTTATCGGGCTC 598
689 TTAGCTTCAAAGAGGCAATCAGATGTAAGTATTCAGACAGCGGTTATCGGGCTC 748
599 CAG-AAGCAGAAATTGCAAAATTCCTTGCCAGGC-TGGCTGCAAGTGTACAGAAATG 656
749 CAGAAACAGAAATTGCAAAATTCCTTGCCAGGC-TGGCTGCAAGTGTACAGAAATG 808
657 TACCT---CAGCTGTGATCTGTGAGCCTAGAAATCATTTTACTGAAATGTTCTCAG 712
809 TACCTCAGCTGTGATCTGTGAGCCTAGAAATCATTTTACTGAAATGTTCTC 868
713 GAATGAATGAAAC 727
869 CAGGAATGAAAAAC 883

RESULT 10
BI259568 875 bp mRNA linear EST 17-JUL-2001
LOCUS BI259568 602968310F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5108094 5',
DEFINITION mRNA sequence.
ACCESSION BI259568
VERSION BI259568.1 GI:14817032
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 875)
AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/BLM at:
<http://image.llnl.gov>
Plate: L14M11261 row: k column: 07
High quality sequence stop: 722.

FEATURES

source

1..875
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="Caxon:9606"
/clone="IMAGE:5108094"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/note="Organ: cervix; Vector: pCMV-Sport6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

ORIGIN

Query Match 47.2%; Score 595; DB 4; Length 875;
Best Local Similarity 91.6%; Pred. No. 7.8e-162;
Matches 809; Conservative 0; Mismatches 50; Indels 24; Gaps 16;

297 CTTTCTCCAAATGTGCGCATCAGCTGTCTGTGTTGAACTCCGTGAGTCAAGTGTTC 356
1 CTTTCTCCAAATGTGCGCATCAGCTGTCTGTGTTGAACTCCGTGAGTCAAGTGTTC 59
357 GGAATTCCTTATTTCCAGTACACAGGTTGTTCCATGATGATATAGATTTGGGCGC 416
60 GGAATTCCTTATTTCCAGTACACAGGTTGTTCCATGATGATATAGATTTGGGCGC 417
417 CCGAGATGTTTGGAGGCCCTTGTCTTTCATCATGAGGCTATGTCAGCGGACT 476
118 CCGAGATGTTT--GGAGGCCCTTGTCTTTCATCATGAGGCTATGTCAGCGGACT 175
477 CAAACGTAACATATGTGAGTGCAGAGATGAAATGTTTAACTCATTTGACTTGG 536
176 CAAACGTAACATATGTGAGTGCAGAGATGAAATGTTTAACTCATTTGACTTGG 233
537 ACTTAGCTTCAAAGAGGCAATCAGATGTAAGTATTCAGACAGCGGTTATCGGGC 596
234 ACTTAGCTTCAAAGAGGCAATCAGATGTAAGTATTCAGACAGCGGTTATCGGGC 293
597 TCCAGAGCAGAAATTGCAAAATTCCTTGCCAGGCCTGCGAGAGTATCAGAAATG 656
294 TCCAGAGCAGAAATTGCAAAATTCCTTGCCAGGCCTGCGAGAGTATCAGAAATG 353
657 TACCTCAGCTGTGATCTGTGAGCCTAGAAATCATTTTACTGAAATGTTCTCAGAAAT 716
354 TACCTCAGCTGTGATCTGTGAGCCTAGAAATCATTTTACTGAAATGTTCTCAGAAAT 413
717 GAAATGAAACATACAGTCAAGTCTCAGAAATGAAAGCAACATTTCTGCTATTTTGA 776
414 GAAATGAAACATACAGTCAAGTCTCAGAAATGAAAGCAACATTTCTGCTATTTTGA 473
777 TCACATATTTGCGAGTAAAGCAGTGTGATTCGCAATTCAGGCTATCAGCTTAAGAGA 836
474 TCACATATTTGCGAGTAAAGCAGTGTGATTCGCAATTCAGGCTATCAGCTTAAGAGA 533
837 CCTTATCAAAAGCAGTGTGATGATTCAGAGAAAGAAATTCCTGCTGAATGGCAAT 896
534 CCTTATCAAAAGCAGTGTGATGATTCAGAGAAAGAAATTCCTGCTGAATGGCA-T 592

QY 897 GTGCACCCATCTTTCAGATTCCTTTTGGCCCCCATATTAAGATCT-GGTGATGCTTC 955
 |||||
 DB 593 GTGCACCCATCTTTCAGATTCCTTTTGGCCCCCATATTAAGATTCGGGTATGCTTC 652
 |||||
 QY 956 CCACTCCAGTG--CTAAGACTGCTGAATGCTGATGATGATTAATCTTGGAGTAA-G 1012
 |||||
 DB 653 CCATCTCAGGGGCTAAAGACTGCTGAATGCTGATGATGATTAATCTTGGAGTAAAG 712
 |||||
 QY 1013 AGGAATATGAAG-ATGTTAGTAAG--ATGTAAAGAGAGTCAAAAATATGACACAG 1069
 |||||
 DB 713 AGGAATATGAAGATGTTAGTGAAGATGAAGAGAGTCAACCAATTTGACCA 772
 |||||
 QY 1070 TGGTATCTACTTGT-----CCAAAGGAAATCTGCGACAGACA-AGTCTTTGTT 1122
 |||||
 DB 773 GGTGATTCCTACTTGTTCCTCAAAAAGAAAATCTGGCGACAGACAGCTTTGTA 832
 |||||
 QY 1123 G-AGTATGCAAT-GCTGTGATTCCAAGCTCGCAGAAATT 1163
 |||||
 DB 833 GAAATATGCAAAATGTTGGGATTCCAAAAGTCCGGAATTT 875
 |||||

RESULT 11
 CF593363 834 bp mRNA linear EST 26-SEP-2003
 LOCUS AGENCOURT_15622279 NIH_MGC_147 Homo sapiens cDNA clone
 DEFINITION IMAGE:30531268 5', mRNA sequence.
 ACCESSION CF593363
 VERSION CF593363.1 GI:36346826
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 JOURNAL 1 (bases 1 to 834)
 COMMENT NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Daniela S. Gerhard, Ph.D.
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: Dr. Stefan Haneson
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
 and advice from Piero Carninci (RIKEN)
 cDNA library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: NDAM621 row: m column: 05
 High quality sequence stop: 660.
 Location/Qualifiers

FEATURES
 source

1. 834
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30531268"
 /issue_type="Human Placenta"
 /lab_host="DH10B TONa"
 /clone_1b="NIH_MGC_147"
 /note="Organ: placenta; Vector: pBluescript; Site 1:
 alt-XhoI; Site 2: BamHI; oligo-dT primed using primer
 5'-TTTTTTTTTTTNN-3', size-selected for average
 insert size 2.3 kb and normalized to 10⁵. This is a
 primary library enriched for full-length clones and
 constructed using the Cap-trapper method (Carninci, in
 preparation). Library constructed by M. Brownstein
 (NIH/NHGRI, National Institutes of Health). Note: This is
 a NIH_MGC library."

Query Match 47.0%; Score 591.6; DB 7; Length 834;
 Best Local Similarity 97.4%; Pred. No. 7.6e-161;
 Matches 644; Conservative 0; Mismatches 10; Indels 7; Gaps 4;

QY 1 ATGGCGGATTCGGGCTGGGCGCTGGGGCGGAGCGCGCGTTTCTGAGGCTTCGGG 60
 |||||
 DB 155 ATGGCGGATTCGGGCTGGGCGCTGGGGCGGAGCGCGCGTTTCTGAGGCTTCGGG 214
 |||||
 QY 61 CGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCTCCGCTCGGTATCGGGTT 120
 |||||
 DB 215 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCTCCGCTCGGTATCGGGTT 274
 |||||
 QY 121 CGCTGTGGCAGCACCCTTGCTGCCCCCGGCGCCCTCAGGCACTTCTTCCGCGAGA 180
 |||||
 DB 275 CGCTGTGGCAGCACCCTTGCTGCCCCCGGCGCCCTCAGGCACTTCTTCCGCGAGA 334
 |||||
 QY 181 ACCACCGGGGCTGGGCGCTCTGGGCGGAGTATGTTTCCGAAAGAGGGCGGGCTG 240
 |||||
 DB 335 ACCACCGGGGCTGGGCGCTCTGGGCGGAGTATGTTTCCGAAAGAGGGCGGGCTG 394
 |||||
 QY 241 GAACAGTTGACAGGTCACAGAAACATGTAATTTGTATGAGTGTTAACAATCCACTT 300
 |||||
 DB 395 GAACAGTTGACAGGTCACAGAAACATGTAATTTGTATGAGTGTTAACAATCCACTT 454
 |||||
 QY 301 TCTCAAAATGTCCATACAGCTGTCTGTCTTGAATCTCTGAGATGTCAGTCTTCGGA 360
 |||||
 DB 455 TCTCAAAATGTCCATACAGCTGTCTGTCTTGAATCTCTGAGATGTCAGTCTTCGGA 514
 |||||
 QY 361 TTGCTCTTATATTCAGTACAGAGGTTGTTTCATGTGATGATACAGATTCGCGCGA 420
 |||||
 DB 515 TTGCTCTTATATTCAGTACAGAGGTTGTTTCATGTGATGATACAGATTCGCGCGA 574
 |||||
 QY 421 GATGTTTGGAGCCCTTCTTCTTCAATCAATGAGGCTATGTCATGCGGACTCAAA 480
 |||||
 DB 575 GATGTTTGGAGCCCTTCTTCTTCAATCAATGAGGCTATGTCATGCGGACTCAAA 634
 |||||
 QY 481 CCACTGATACATTTGTGAGTGCAGAGATGAATGTTTAACTATGACTTGGACTT 540
 |||||
 DB 635 CCACTGATACATTTGTGAGTGCAGAGATGAATGTTTAACTATGACTTGGACTT 694
 |||||
 QY 541 A-GCTTCAAGAAAGCAATCAGATGTAAAGTATTCAGACAGCGGCTATC-GGGCTC 598
 |||||
 DB 695 AGGCTTCAAGAAAGCAATCAGATGTAAAGTATTCAGACAGCGGCTATC-GGGCTC 754
 |||||
 QY 599 CAGAAGCAGAAATTCAGAAATTCCTT---GGCCAGGCT--GGCTGCAGAGTATACAGA 653
 |||||
 DB 755 CAGAAGCAGAAATTCAGAAATTCCTTGGGCCCCAGGCTTGCCCTGCAGAGATACAGA 814
 |||||
 QY 654 A 654
 |||||
 DB 815 A 815
 |||||

RESULT 12
 BG386905 906 bp mRNA linear EST 12-MAR-2001
 LOCUS 602454704F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4583065 5',
 DEFINITION mRNA sequence.
 ACCESSION BG386905
 VERSION BG386905.1 GI:13280454
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 JOURNAL 1 (bases 1 to 906)
 COMMENT NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: NIH Intramural Sequencing Center
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 plate: LNCM1306 row: 9 column: 02
 High quality sequence stop: 668.

FEATURES

source

1..906
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4583065"
 /issue_type="adenocarcinoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_15"

/note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; CDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGACCGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using Zap-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)"

ORIGIN

Query Match 46.3%; Score 583.8; DB 4; Length 906;
 Best Local Similarity 96.1%; Pred. No. 1.5e-158;
 Matches 664; Conservative 0; Mismatches 17; Indels 10; Gaps 6;

```

QY 570 GATATTTCAGACAGAGGGGATCGGGCTCCAGAGCAGAAATTCGTTGGCCCA 629
DB 2 GATATTTCAGACAGAGGGGATCGGGCTCCAGAGCAGAAATTCGTTGGCCCA 61
QY 630 GGGTCGCTGCAGAGTATCAGAAATGTAAGTCTGATCTGTGAGCCTTAGAAT 689
DB 62 GGGTCGCTGCAGAGTATCAGAAATGTAAGTCTGATCTGTGAGCCTTAGAAT 121
QY 690 CATTTTACGGAATGTTCTCAGGAATGAAACATACAGTCAATCTCAGGAATG 749
DB 122 CATTTTACGGAATGTTCTCAGGAATGAAACATACAGTCAATCTCAGGAATG 181
QY 750 GAAGGCAACAGTTCGTATTTATGATCATATTTGGCAGTAAGAGTGTGAATG 809
DB 182 GAAGGCAACAGTTCGTATTTATGATCATATTTGGCAGTAAGAGTGTGAATG 241
QY 810 CGCAATTCAGCCTTATCAGTAAGAGCCTTATCAAAAGCATCTTATGATGATCA 869
DB 242 CGCAATTCAGCCTTATCAGTAAGAGCCTTATCAAAAGCATCTTATGATGATCA 301
QY 870 CAGAGAAATTCCTGCTGAAATGCGATTTGCGACCCATCTTTAGCATCTTTGCCC 929
DB 302 CAGAGAAATTCCTGCTGAAATGCGATTTGCGACCCATCTTTAGCATCTTTGCCC 361
QY 930 TCATATTGAAGATCTGTCAATGCTTCCACTCCAGTCTAAGACCTGTAAGTGTGGA 989
DB 362 TCATATTGAAGATCTGTCAATGCTTCCACTCCAGTCTAAGACCTGTAAGTGTGGA 421
QY 990 TGAATGATATCTTGGAAATGAAGAAATGAAGATGTTGAAGATGTAAGAAGGA 1049
DB 422 TGAATGATATCTTGGAAATGAAGAAATGAAGATGTTGAAGATGTAAGAAGGA 481
QY 1050 GTGTCAAAAATATGACCAAGTGTATCTCTAATCTTTCCAAAGAAAATCTGGCAGAG 1109
DB 482 GTGTCAAAAATATGACCAAGTGTATCTCTAATCTTTCCAAAGAAAATCTGGCAGAG 539
QY 1110 ACAAGCTTTGTTGAGATGCAATGCTGTGATTCGAAAGCTGCGAGAAATTAATGAC 1169
DB 540 ACAAGCTTTGTTGAGATGCAATGCTGTGATTCGAAAGCTGCGAGAAATTAATGAC 598
QY 1170 TGAAGAGATTTGATGGAAGATTTGTTGTGCTACATTTACCCGCTGAGTCCATCA 1229
DB 599 TGAAGAG-ATGTTGATGGAAGATTTGTGT---GGTACATTTACCCGCTGAGTCC- -TACA 651

```

QY 1230 GAGGGATATCTGTATCAACCTTGCTTAA 1260
 DB 652 AAGAGGATATCTGTATCAACCTTGCTTAA 682

RESULT 13
 B1333581
 LOCUS
 DEFINITION
 mRNA sequence.
 B1333581
 ACCESSION
 B1333581
 VERSION
 B1333581.1 GI:15018238
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 NIH-MGC <http://mgs.nci.nih.gov/>.
 1 (bases 1 to 682)
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC

CDNA library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 plate: LNCM1342 row: 3 column: 17
 High quality sequence stop: 674.
 Location/Qualifiers

FEATURES

source

1..682
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5139184"
 /issue_type="cervical carcinoma cell line"
 /lab_host="DH10B"

/note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.4 kb. Library prepared by Life
 Technologies."

ORIGIN

Query Match 46.3%; Score 583.6; DB 4; Length 682;
 Best Local Similarity 97.1%; Pred. No. 1.5e-158;
 Matches 647; Conservative 0; Mismatches 14; Indels 5; Gaps 5;

```

QY 1 ATGGCGGATCCGGCTGCGCTGCGGCGGAGCGCGCGCTTTCTGAGGCTTCGGG 60
DB 19 ATGGCGGATCCGGCTGCGCTGCGGCGGAGCGCGCGCTTTCTGAGGCTTCGGG 78
QY 61 CGGCTGTGCGAGTACAGAGCGCTCTGGGTAGCGGCTCTCGGCTCGGTATGAGGTT 120
DB 79 CGGCTGTGCGAGTACAGAGCGCTCTGGGTAGCGGCTCTCGGCTCGGTATGAGGTT 138
QY 121 CGGTGTGCGGAGACCTGAGCTCGCCCGCGGCGCGCTCAGAGCATCTTGGCGCAGGA 180
DB 139 CGGTGTGCGGAGACCTGAGCTCGCCCGCGGCGCGCTCAGAGCATCTTGGCGCAGGA 198
QY 181 ACCACCGGGGCTGCGGCTCTGCGCGGAGTATGTTTCCGAAAGAGAGGCGGCGCTG 240
DB 199 ACCACCGGGGCTGCGGCTCTGCGCGGAGTATGTTTCCGAAAGAGAGGCGGCGCTG 258
QY 241 GAACAGTTGCAAGGTCACAGAAACATCGTACTTTGTATGAGAGTGTTCATCACTTT 300
DB 259 GAACAGTTGCAAGGTCACAGAAACATCGTACTTTGTATGAGAGTGTTCATCACTTT 318
QY 301 TTCGAAATGTGCATCAGCTGTCTGTGTTGTAAGTCTGTGATGTCAGTGTTTCGGA 360

```

```

Db      319  TCCCAAAATGTCATACAGCTGTCTGTGCTTGAAGTCTCTGGATGTCAGTGTTCGAA 378
Qy      361  TTGCTCTTTATATCCAGTCAACAGGGTGTTCATGTGATGATGATACAGATTTGCGCCGA 420
Db      379  TTGCTCTTTATATCCAGTCAACAGGGTGTTCATGTGATGATGATACAGATTTGCGCCGA 438
Qy      421  GATGTTTGGAGGCCCTTGTCTTTTCATCATGAGGCTATGTCCATGCGAGCTCAAA 480
Db      439  GATG-TTGGAGGCCCTTGTCTTTTCATCATGAGGCTATGTCCATGCGAGCTCAAA 497
Qy      481  CCAAGTAACTATTTGGAGTGCAGAGATGAATGTTTAACTCATGACTTTGGAATT 540
Db      498  CCAAGTAACTATTTGGAGTGCAGAGATGAATGTTTAACTCATGACTTTGGAATT 557
Qy      541  AGCTTAAAGAGGCAATCA-AGATGTAAGTATATTCACAGAG-ACGGGTATCGGGCTC 598
Db      558  AGCTTAAAGAGGCAATCAAGATGTAAGTATATTCACAGAGACGAGGTATCGGGCTC 617
Qy      599  CAGAGCAGATTTGCAAAATTTGCTTGGCCAGGCTGCGCTGACAGATGATA-CAGATGT 657
Db      618  CAGAGCAGATTTGCAAAATTTGCTTGG-CCAGGCTGCGCTGACAGATGATACAGATGT 676
Qy      658  ACCTCA 663
Db      677  ACCTCA 682

```

RESULT 14
BQ225071 888 bp mRNA linear EST 02-MAY-2002
LOCUS AGENCOURT 7323233 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:6016796
DEFINITION 5', mRNA Sequence.

ACCESSION BQ225071
VERSION BQ225071.1 GI:20406471
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE NIH-MGC http://mgc.nci.nih.gov/
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LIML)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LIML at:
http://image.liml.gov
Plate: LIML3215 row: a column: 21
High quality sequence stop: 581.

FEATURES
source
Location/Qualifiers

1..888
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6016796"
/tissue_type="epithelioid carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH MGC 70"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.1 kb. Library constructed by Life
Technologies."

ORIGIN

Query Match 45.9%; Score 578; DB 5; Length 888;
Best Local Similarity 97.8%; Pred. No. 7.2e-157;
Matches 618; Conservative 0; Mismatches 10; Indels 4; Gaps 3;

```

Qy      1  ATGCGCGGATCCGGCTGCGCTGCGGCGCGAGCGCGCGCTTTCTGAGGCTTTCGGG 60
Db      149  ATGCGCGGATCCGGCTGCGCTGCGGCGCGAGCGCGCGCTTTCTGAGGCTTTCGGG 208
Qy      61  CGGCTGTGGCAGATACAGAGCGCTCTGGGTACCGCTCTCGGCTCGGTATAGGGTT 120
Db      209  CGGCTGTGGCAGATACAGAGCGCTCTGGGTACCGCTCTCGGCTCGGTATAGGGTT 268
Qy      121  CGGCTGTGGGGAACCTTGGCTCGCCCCCGCGCGCTTCAGGAGTTCTTTCGCGCAGA 180
Db      269  CGGCTGTGGGGAACCTTGGCTCGCCCCCGCGCGCTTCAGGAGTTCTTTCGCGCAGA 328
Qy      181  ACCACCGGGGCTGCGGCTCTGCGCGCGAGTATGTTTCCGAAAGAGAGGCGGCGCTG 240
Db      329  ACCACCGGGGCTGCGGCTCTGCGCGCGAGTATGTTTCCGAAAGAGAGGCGGCGCTG 388
Qy      241  GAACAGTTGCAGGCTCACAGAAACATCGTACTTTGTATGAGTGTATTCACACTTT 300
Db      389  GAACAGTTGCAGGCTCACAGAAACATCGTACTTTGTATGAGTGTATTCACACTTT 448
Qy      301  TCTCCAAATGTGCATCAGCTGTCTGTGCTTGAATCTCGATGTCACTGTTCGGA 360
Db      449  TCTCCAAATGTGCATCAGCTGTCTGTGCTTGAATCTCGATGTCACTGTTCGGA 508
Qy      361  TTGCTCTTTATATTCAGTCAACGAGGTTGTTTCATGTGATGATACAGATTTGCGCCGA 420
Db      509  TTGCTCTTTATATTCAGTCAACGAGGTTGTTTCATGTGATGATACAGATTTGCGCCGA 568
Qy      421  GATGTTTGGAGG-CCCTTGTCTTTTCTTATCATGAGGCTATGTCCATGCGGACTCAA 479
Db      569  GATGTTTGGAGGCGCCCTTGTCTTTTCTTATCATGAGGCTATGTCCATGCGGACTCAA 628
Qy      480  ACCACGTAACATATTGTGAGTGCAGAGATGAATGTTTAACTCATGACTTTGGAAT 539
Db      629  ACCACGTAACATATTGTGAGTGCAGAGATGAATGTTTAACTCATGACTTTGGAAT 688
Qy      540  TAGCTTCAAGAA-GGCAATCAGATGTAAGTATATTCAGACAGCGGT--ATCGGCT 596
Db      689  TAGCTTCAAGAAAGGAATCAAGATGTAAGTATATTCAGACGAGCGGTATCGGGCT 748
Qy      597  TCCAGAGCGAATTTGCCAAATTTGCTTGGCCC 628
Db      749  TCCAGAGCGAATTTGCCAAATTTGCTTGGCCC 780

```

RESULT 15

LOCUS COJ37403 602 bp mRNA linear EST 29-JUL-2004
DEFINITION S1LT02c17116f1 squirrel testis library 1 Spermophilus lateralis
CDNA clone 17116 5', mRNA sequence.

ACCESSION COJ37403
VERSION COJ37403.1 GI:50824673
KEYWORDS EST.

SOURCE Spermophilus lateralis (golden-mantled ground squirrel)
ORGANISM Spermophilus lateralis

REFERENCE 1 (bases 1 to 602)
AUTHORS Williams,D.R., Gracey,A.Y., Martin,S.L., Hughes,M.A., Li,W.,
Rogers,J. and Cossins,A.R.
TITLE Microarray analysis of transcriptional changes during hibernation
COMMENT in the golden mantled ground squirrel, Spermophilus lateralis
Unpublished (2004)
Contact: Andrew R. Cossins
Laboratory for Environmental Gene Regulation
University of Liverpool
School of Biological Sciences, The Biosciences Building, Crown
Street, Liverpool, United Kingdom, L69 7ZB
Tel: +44(0)151-795-4510
Fax: +44(0)151-795-4531
Email: cossins@liv.ac.uk

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2004, 11:23:30 ; Search time 5510 Seconds
(without alignments)
10813.978 Million cell updates/sec

Title: US-10-798-532-3

Perfect score: 1260

Sequence: 1 atggcggggagtcgagctgcgc.....tgatcaaacctgtcttaa 1260

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapect 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sbs:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1255.2	99.6	1260	6	CO715025
2	1255.2	99.6	1260	6	AX262521
3	1255.2	99.6	1258	6	AX262519
4	1255.2	99.6	2909	9	HSAS36197
5	1255.2	99.6	4065	6	CO412057
6	1255.2	99.6	4065	6	AX188296
7	1253.6	99.5	1260	6	AX166543
8	1253.6	99.5	1824	6	AX680136
9	1114.4	88.2	2041	10	BC058732
10	1111.2	88.2	110000	2	AC129141_1
11	1106.4	87.8	1633	10	RNR1SRNA
12	1106.4	87.8	3244	10	RNT70372
13	1098.4	87.2	1254	10	MMPKXIS
14	1065.6	84.6	1357	9	BC026046
15	989.8	78.6	1135	9	AK058195
16	941.8	74.7	2041	9	BC014917
17	768	61.0	1157	5	CR386837
18	699.8	55.5	2940	9	HSMB05327
19	567.6	45.0	259562	2	AC094499

20	535.4	42.5	601	10	MMK1SRNA	X82320 M.musculus
21	458.6	36.4	529	6	AX186343	AX186343 Sequence
22	445.8	35.4	722	6	BD221413	BD221413 Human gen
23	426.6	33.9	470	6	AX187546	AX187546 Sequence
24	387.4	30.7	462	6	AX184680	AX184680 Sequence
25	295	23.4	105902	9	ALJ59659	ALJ59659 Human DNA
26	295	23.4	167995	2	AC040901	AC040901 Homo sapi
27	293.4	23.3	211812	2	AC027205	AC027205 Homo sapi
28	254.4	20.2	200095	10	AC121606	AC121606 Mus muscu
29	254.4	20.2	220806	10	AC123650	AC123650 Mus muscu
30	252.8	20.1	100207	10	AY180177	AY180177 Mus muscu
31	249.6	19.8	211635	2	AC112763	AC112763 Rattus no
32	203.8	16.2	231	6	BD220137	BD220137 Human gen
33	184.8	14.7	185810	2	AC136052	AC136052 Rattus no
34	177.4	14.1	419	6	CQ393016	CQ393016 Sequence
35	177.4	14.1	419	6	CQ399411	CQ399411 Sequence
36	164.6	13.1	60640	10	AY180176	AY180176 Mus muscu
37	153.4	12.2	205752	2	AC119431	AC119431 Mus muscu
38	87.4	6.9	185810	2	AC136052	AC136052 Rattus no
39	85	6.7	198	6	CQ405796	CQ405796 Sequence
40	69.6	5.5	2108	8	NEL18351	Y18351 Nicotiana p
41	64.6	5.1	2047	8	BT013617	BT013617 Lycopersi
42	63.4	5.0	65	6	CQ555763	CQ555763 Sequence
43	60.6	4.8	1996	5	BC044032	BC044032 Xenopus 1
44	59.8	4.7	2204	8	NPL18350	Y18350 Nicotiana p
45	55.8	4.4	1565	5	BC065869	BC065869 Danto rer

ALIGNMENTS

RESULT 1	CO715025	Sequence 959 from Patent WO02068579.	1260 bp	DNA	linear	PAT 03-FEB-2004
LOCUS	CO715025					
DEFINITION	Sequence 959 from Patent WO02068579.					
ACCESSION	CO715025					
VERSION	CO715025.1	GI:42275882				
KEYWORDS						
SOURCE						
ORGANISM	Homo sapiens (human)					
REFERENCE						
AUTHORS	Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.					
TITLE	Kite, such as nucleic acid arrays, comprising a majority of humenexons or transcripts, for detecting expression and other uses thereof					
JOURNAL	Patent: WO 02068579-A 959 06-SBP-2002; PE Corporation (NY) (US)					
FEATURES						
source	1..1260	Location/Qualifiers				
ORIGIN		/organism="Homo sapiens"				
		/mol_type="unassigned DNA"				
		/db_xref="taxon:9606"				
Query Match	99.6%;	Score 1255.2;	DB 6;	Length 1260;		
Best local Similarity	99.8%;	Pred. No. 1.2e-310;				
Matches 1257;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;		
QY	1	ATGGCGGAGTCCGGCTGAGCGCTGAGCGCGCGCGCTTTCTGAGGCGCTTCGGG	60			
DB	1	ATGGCGGAGTCCGGCTGAGCGCTGAGCGCGCGCGCTTTCTGAGGCGCTTCGGG	60			
QY	61	CGGCTGTGACAGTACAGAGCGGTCTGGGTAGCGGCTCTCGGCTCGGTATCGGGTT	120			
DB	61	CGGCTGTGACAGTACAGAGCGGTCTGGGTAGCGGCTCTCGGCTCGGTATCGGGTT	120			
QY	121	CGGTGTGCGGCAACCTTGCTGCGCCCGCGCGCGCTTACGACATGTTCTTCCGCCAAGA	180			
DB	121	CGGTGTGCGGCAACCTTGCTGCGCCCGCGCGCGCTTACGACATGTTCTTCCGCCAAGA	180			
QY	181	ACACCGGGGCTGCGGCTCTGCGCGCGAGTATGTTTCCGAAAGAGCGCGCGCTG	240			

```
DB 181 ACCACGGGGGCTGGCCCTCTGCGCGAGTATGTTTCCGCAAGAGAGGGGCGCGCTG 240
QY 241 GAACAGTTGAGGGGTACAGAAACATCGTACTTTGTATGATGTTTCAATCCACTTT 300
DB 241 GAACAGTTGAGGGGTACAGAAACATCGTACTTTGTATGATGTTTCAATCCACTTT 300
QY 301 TCTCCAAATGTGCGCATCAGCGTGTCTGTGCTTGAACCTCGTGAGTCAAGTGTTCGAA 360
DB 301 TCTCCAAATGTGCGCATCAGCGTGTCTGTGCTTGAACCTCGTGAGTCAAGTGTTCGAA 360
QY 361 TTGCTCTTATATTCAGTCAACCGGTTGTTCATGTGATGATTAAGCATTTGGCGCGCA 420
DB 361 TTGCTCTTATATTCAGTCAACCGGTTGTTCATGTGATGATTAAGCATTTGGCGCGCA 420
QY 421 GATGTTTGGAGGCCCTTGCTTTCTTCATCATGAGGCGTATGCTCAATGCGGACCTCAAA 480
DB 421 GATGTTTGGAGGCCCTTGCTTTCTTCATCATGAGGCGTATGCTCAATGCGGACCTCAAA 480
QY 481 CCAAGTAAATATTTGTGAGTGCAGAGATGAATGTTTAACTCATATGACTTTGAGACTT 540
DB 481 CCAAGTAAATATTTGTGAGTGCAGAGATGAATGTTTAACTCATATGACTTTGAGACTT 540
QY 541 AGCTTCAAGAGAGGCATCAGATGTAAGTATATTCAGACAGCGGTTATCGGGCTCCA 600
DB 541 AGCTTCAAGAGAGGCATCAGATGTAAGTATATTCAGACAGCGGTTATCGGGCTCCA 600
QY 601 GAAGCAGAAATTGCAAAATTTGCTGGCCAGGCTGCGCGTCAAGATGATCAGAAATGACC 660
DB 601 GAAGCAGAAATTGCAAAATTTGCTGGCCAGGCTGCGCGTCAAGATGATCAGAAATGACC 660
QY 661 TCAGCTGTGATCTGTGAGGCTTGAAGTATCTGAAATGTTTCTCAGGAATGAAA 720
DB 661 TCAGCTGTGATCTGTGAGGCTTGAAGTATCTGAAATGTTTCTCAGGAATGAAA 720
QY 721 CTGAAACATTCAGTCAAGTCTCAGAAATGAGAGCAACGTTCTGCTATATTTGATCAC 780
DB 721 CTGAAACATTCAGTCAAGTCTCAGAAATGAGAGCAACGTTCTGCTATATTTGATCAC 780
QY 781 ATATTTGGCCAGTAAAGAGTGAATGCGCAATTCAGCTATCACTTAAGAGACCTT 840
DB 781 ATATTTGGCCAGTAAAGAGTGAATGCGCAATTCAGCTATCACTTAAGAGACCTT 840
QY 841 ATCAAAAGCATGCTTCATGATGATCCAGCAGAAATTCCTGCTGAATGAGCATTTGTC 900
DB 841 ATCAAAAGCATGCTTCATGATGATCCAGCAGAAATTCCTGCTGAATGAGCATTTGTC 900
QY 901 AGCCCATTTCTTATGATCTTTTGGCCCTCATATTTGAAGATCGTGCATGCTTCCACT 960
DB 901 AGCCCATTTCTTATGATCTTTTGGCCCTCATATTTGAAGATCGTGCATGCTTCCACT 960
QY 961 CCAAGTCTAAGTCAAGTCAAGTGAATGCGGAGATGATATTTGGGAATGAAGAGAAAT 1020
DB 961 CCAAGTCTAAGTCAAGTCAAGTGAATGCGGAGATGATATTTGGGAATGAAGAGAAAT 1020
QY 1021 GAAGATGTTGTAGAAGATGTAAGAGAGAGATGTAAGAGATGTAAGAGATGTAAGAT 1080
DB 1021 GAAGATGTTGTAGAAGATGTAAGAGAGAGATGTAAGAGATGTAAGAGATGTAAGAT 1080
QY 1081 CTTGTTTCAAAAGAAAATCTTGCGCAGAGACAAGTCTTTTGTGAGTATGCAATGCTGCT 1140
DB 1081 CTTGTTTCAAAAGAAAATCTTGCGCAGAGACAAGTCTTTTGTGAGTATGCAATGCTGCT 1140
QY 1141 GATTTCCAAAGCTCGCGAATTAATGATCAGAGAGATGTTTGAATGAGAAATTTGTTG 1200
DB 1141 GATTTCCAAAGCTCGCGAATTAATGATCAGAGAGATGTTTGAATGAGAAATTTGTTG 1200
QY 1201 GCTACATTTCAACCGCTGAGTGCCTACAGAGGGGATATCTGTATCAAACTTGCTTTAA 1260
DB 1201 GCTACATTTCAACCGCTGAGTGCCTACAGAGGGGATATCTGTATCAAACTTGCTTTAA 1260
```

```
AX262521
LOCUS AX262521 1260 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 12 from Patent WO0173050.
ACCESSION AX262521
VERSION AX262521.1 GI:16511408
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Meyers, R.
AUTHORS 3714, 16742, 23546, and 13887 novel protein kinase molecules and
TITLE uses therefor
JOURNAL Patent: WO 0173050-A 12 04-OCT-2001;
FEATURES Millennium Pharmaceuticals, Inc. (US)
source 1. 1260
Location/Qualifiers
1. 1260
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.6%; Score 1255.2; DB 6; Length 1260;
Best Local Similarity 99.8%; Pred. No. 1.2e-310; Mismatches 0; Gaps 0;
Matches 1257; Conservative 0; Indels 3;
QY 1 ATGGCGGAGATCCGGCTGCGCTGCGGCGGAGCGCGCGCTTTCTGAGGCGCTTCGGG 60
DB 1 ATGGCGGAGATCCGGCTGCGCTGCGGCGGAGCGCGCGCGCTTTCTGAGGCGCTTCGGG 60
QY 61 CGGCTGTGCGAGTACAGAGCCGTCTGAGTACCGGCTCTCGGCTCGGTATGCGGTT 120
DB 61 CGGCTGTGCGAGTACAGAGCCGTCTGAGTACCGGCTCTCGGCTCGGTATGCGGTT 120
QY 121 CGCTGTGCGGACACCTTGCGCTGCGCCCGGCGGCGCTTTCAGGACATTTGCGCGAGA 180
DB 121 CGCTGTGCGGACACCTTGCGCTGCGCCCGGCGGCGCTTTCAGGACATTTGCGCGAGA 180
QY 181 ACCACCGGGGCTGCGGCTCTGCGCCCGAGTATGTTTCCGAAAGAGAGGCGCGCTG 240
DB 181 ACCACCGGGGCTGCGGCTCTGCGCCCGAGTATGTTTCCGAAAGAGAGGCGCGCTG 240
QY 241 GAACAGTTGAGGGGTACAGAAACATCGTACTTTGTATGATGTTTCAATCCACTTT 300
DB 241 GAACAGTTGAGGGGTACAGAAACATCGTACTTTGTATGATGTTTCAATCCACTTT 300
QY 301 TCTCCAAATGTGCGCATCAGCGTGTCTGTGCTTGAACCTCGTGAGTCAAGTGTTCGAA 360
DB 301 TCTCCAAATGTGCGCATCAGCGTGTCTGTGCTTGAACCTCGTGAGTCAAGTGTTCGAA 360
QY 361 TTGCTCTTATATTCAGTCAACCGGTTGTTCATGTGATGATTAAGCATTTGGCGCGCA 420
DB 361 TTGCTCTTATATTCAGTCAACCGGTTGTTCATGTGATGATTAAGCATTTGGCGCGCA 420
QY 421 GATGTTTGGAGGCCCTTGCTTTCTTCATCATGAGGCGTATGCTCAATGCGGACCTCAAA 480
DB 421 GATGTTTGGAGGCCCTTGCTTTCTTCATCATGAGGCGTATGCTCAATGCGGACCTCAAA 480
QY 481 CCAAGTAAATATTTGTGAGTGCAGAGATGAATGTTTAACTCATATGACTTTGAGACTT 540
DB 481 CCAAGTAAATATTTGTGAGTGCAGAGATGAATGTTTAACTCATATGACTTTGAGACTT 540
QY 541 AGCTTCAAGAGAGGCATCAGATGTAAGTATATTCAGACAGCGGTTATCGGGCTCCA 600
DB 541 AGCTTCAAGAGAGGCATCAGATGTAAGTATATTCAGACAGCGGTTATCGGGCTCCA 600
QY 601 GAAGCAGAAATTGCAAAATTTGCTGGCCAGGCTGCGCGTCAAGATGATCAGAAATGACC 660
DB 601 GAAGCAGAAATTGCAAAATTTGCTGGCCAGGCTGCGCGTCAAGATGATCAGAAATGACC 660
QY 661 TCAGCTGTGATCTGTGAGGCTTGAAGTATCTGAAATGTTTCTCAGGAATGAAA 720
DB 661 TCAGCTGTGATCTGTGAGGCTTGAAGTATCTGAAATGTTTCTCAGGAATGAAA 720
```

```

Db      661 TCAGCTGTTGATCTGTGAGCCCTAAGATCATTTTATCTGGAATGTTCTCAGGAATGAAA 720
Qy      721 CTGAACATACATGATCTGAGATGGAATGGAAGCAAAAGTTCTGCTATTTATGATGAC 780
Db      721 CTGAACATACATGATCTGAGATGGAATGGAAGCAAAAGTTCTGCTATTTATGATGAC 780
Qy      781 ATATTTGCCAGTAAAGCAGTGTGAAATGCCGCAATTCAGCCTATGACCTTAAGAGACCTT 840
Db      781 ATATTTGCCAGTAAAGCAGTGTGAAATGCCGCAATTCAGCCTATGACCTTAAGAGACCTT 840
Qy      841 ATCAAAAGCATGCTTATGATGATCCAGACAGAAATTCCTGCTGAATGCGATTGTGC 900
Db      841 ATCAAAAGCATGCTTATGATGATCCAGACAGAAATTCCTGCTGAATGCGATTGTGC 900
Qy      901 AGCCCATTTCTTGAAGATTCCTTTGCCCCCTCATATTTGAAGATGCTGGCAATGCTCCACT 960
Db      901 AGCCCATTTCTTGAAGATTCCTTTGCCCCCTCATATTTGAAGATGCTGGCAATGCTCCACT 960
Qy      961 CCAGTCTAAGACTGCTGATGTGCTGATGATGATTTATCTTGGGAATGAAGAGAAATAT 1020
Db      961 CCAGTCTAAGACTGCTGATGTGCTGATGATGATTTATCTTGGGAATGAAGAGAAATAT 1020
Qy      1021 GAAGATGTTGTAAGATGTAATAAGAGAGTCAAAATATGACCAAGTGTATCTCTA 1080
Db      1021 GAAGATGTTGTAAGATGTAATAAGAGAGTCAAAATATGACCAAGTGTATCTCTA 1080
Qy      1081 CTGTGTCGAAGAAATCCTGGCAGAGCAAGTCTTGTGTAAGTGAATGAATGCTGGT 1140
Db      1081 CTGTGTCGAAGAAATCCTGGCAGAGCAAGTCTTGTGTAAGTGAATGAATGCTGGT 1140
Qy      1141 GATTCCAAAGCTCGCAGAAATTAATGACTGAGAGATGTTGATGGAAAGTTGTTGTG 1200
Db      1141 GATTCCAAAGCTCGCAGAAATTAATGACTGAGAGATGTTGATGGAAAGTTGTTGTG 1200
Qy      1201 GCTACATTTCTACCCGCTGAGTGCCTAACAAGAGGAGATCTGATCAAACTTGTCTTAA 1260
Db      1201 GCTACATTTCTACCCGCTGAGTGCCTAACAAGAGGAGATCTGATCAAACTTGTCTTAA 1260

```

```

RESULT 3
AX262519      2598 bp      DNA      linear      PAT 26-OCT-2001
LOCUS      AX262519
DEFINITION      Sequence 10 from Patent WO0173050.
ACCESSION      AX262519
VERSION      AX262519.1 GI:16511406
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS      Meyers, R.
TITLE      3714, 16742, 23546, and 13887 novel protein kinase molecules and
JOURNAL      uses thereof
Patent: WO 0173050-A 10 04-OCT-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
Source
1. 2598
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
268..1527
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD10238.1"
/db_xref="GI:16511407"
/translation="MAGSGCAGMGAEPRLFARGLMVOYSLGSGSASVYVRCCG
NPGSPGALKOPLPGTTGAASAAYGPKREBALBLOQGHNTITVIGVTHFSP
NYSRCLILELDVYSBELIYSHOGCSKMTQHKARDYLEALILHBYHDLK
PRLIWSAENECFKLIDFGLSFKENQDVYTIQTDI RAEVALQCLQVAGQSTDE
CTSAVDLWSIGILILEWFSGMKLKHTVRSQEMKANSADIDHIFAKAVNNAIPYH
LRILKSMLEHDDPSRIIPAMALCSPPFSIPFAPHIEDLVMLPTPIRLINLVDDYL
ENEEYEDVDVDEKQKYGIVSVLLVRENPRGQVFVEVYNAAGDSKAAQLTGR

```

```

ORIGIN
MEDKGVATFVPLSAVRYLYOTLL"
Query Match      99.6%; Score 1255.2; DB 6; Length 2598;
Best Local Similarity 99.8%; Pred. No. 1.4e-310;
Matches 1257; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy      1 ATGGCGGATCCGGCTGCGCTGCGGCGCGAGCCGCGCTTTCTGAGAGCCTTCGGG 60
Db      268 ATGGCGGATCCGGCTGCGCTGCGGCGCGAGCCGCGCTTTCTGAGAGCCTTCGGG 327
Qy      61 CGGCTGTGGCAGTACAGAGCCCTTGGGTAGCCGCTCTCCGCTCGGTATGCGGT 120
Db      328 CGGCTGTGGCAGTACAGAGCCCTTGGGTAGCCGCTCTCCGCTCGGTATGCGGT 387
Qy      121 CGGTGTGGGCAACCTGGGCTGCGGCCCGCGGCGCTCAGGCAATGCTTGGCGCAGGA 180
Db      388 CGGTGTGGGCAACCTGGGCTGCGGCCCGCGGCGCTCAGGCAATGCTTGGCGCAGGA 447
Qy      181 ACCACCGGGGCTGCGGCTGCTGCGCGGAGTATGCTTCCGAAAGAGAGGCGCGCTG 240
Db      448 ACCACCGGGGCTGCGGCTGCTGCGCGGAGTATGCTTCCGAAAGAGAGGCGCGCTG 507
Qy      241 GAACAGTTGACAGGTTCAGAGAAACATCGTGACTTGTATGAGTGTTCATTCACCTT 300
Db      508 GAACAGTTGACAGGTTCAGAGAAACATCGTGACTTGTATGAGTGTTCATTCACCTT 567
Qy      301 TCTCCAAATGTCCTACAGCTGTCTGTGCTTGAACCTCGGATGTCAGTGTTCGGAA 360
Db      568 TCTCCAAATGTCCTACAGCTGTCTGTGCTTGAACCTCGGATGTCAGTGTTCGGAA 627
Qy      361 TTGCTCTTATTTCCAGTCAACAGAGGTTGTTCCATGGATGATACAGATTCGCCGA 420
Db      628 TTGCTCTTATTTCCAGTCAACAGAGGTTGTTCCATGGATGATACAGATTCGCCGA 687
Qy      421 GATGTTTGGAGCCCTTGTCTTCTTCAATCAATGAGGCTATGTCATGCGGACTCAAA 480
Db      688 GATGTTTGGAGCCCTTGTCTTCTTCAATCAATGAGGCTATGTCATGCGGACTCAAA 747
Qy      481 CCAGCTAATATTTGTGAGTGCAGAGATGAATGTTTAACTCATTTGATCTTGAAGCTT 540
Db      748 CCAGCTAATATTTGTGAGTGCAGAGATGAATGTTTAACTCATTTGATCTTGAAGCTT 807
Qy      541 AGCTTCAAGAGGCAATCAGAGATGTAAGTATTTACAGACAGAGGATGCGGCTCCA 600
Db      808 AGCTTCAAGAGGCAATCAGAGATGTAAGTATTTACAGACAGAGGATGCGGCTCCA 867
Qy      601 GAAGCAATTCGAAATTTGCTTGGCCGAGGCTGCGCTGAGATGATACAGATGTATC 660
Db      868 GAAGCAATTCGAAATTTGCTTGGCCGAGGCTGCGCTGAGATGATACAGATGTATC 927
Qy      661 TCAGCTGTTGATCTGTGAGCCCTAAGATCATTTTATCTGGAATGTTCTCAGGAATGAAA 720
Db      928 TCAGCTGTTGATCTGTGAGCCCTAAGATCATTTTATCTGGAATGTTCTCAGGAATGAAA 987
Qy      721 CTGAACATACATGATCTGAGATGGAATGGAAGCAAAAGTTCTGCTATTTATGATGAC 780
Db      988 CTGAACATACATGATCTGAGATGGAATGGAAGCAAAAGTTCTGCTATTTATGATGAC 1047
Qy      781 ATATTTGCCAGTAAAGCAGTGTGAAATGCCGCAATTCAGCCTATGACCTTAAGAGACCTT 840
Db      1048 ATATTTGCCAGTAAAGCAGTGTGAAATGCCGCAATTCAGCCTATGACCTTAAGAGACCTT 1107
Qy      841 ATCAAAAGCATGCTTATGATGATCCAGACAGAAATTCCTGCTGAATGCGATTGTGC 900
Db      1108 ATCAAAAGCATGCTTATGATGATCCAGACAGAAATTCCTGCTGAATGCGATTGTGC 1167
Qy      901 AGCCCATTTCTTGAAGATTCCTTTGCCCCCTCATATTTGAAGATGCTGGCAATGCTCCACT 960
Db      1168 AGCCCATTTCTTGAAGATTCCTTTGCCCCCTCATATTTGAAGATGCTGGCAATGCTCCACT 1227
Qy      961 CCAGTCTAAGACTGCTGATGTGCTGATGATGATTTATCTTGGGAATGAAGAGAAATAT 1020

```

Db 1228 CCAGTGTCTAGACTGCTGTAATGTGCTGATGATGATTAATCTTGAGATGAGAGGAATAT 1287
QY 1021 GAAGATGTTGTAGAAAGATGTAAGAGAGAGAGTGTCAAAAATATGACACAGTGTATCTTA 1080
Db 1288 GAAGATGTTGTAGAAAGATGTAAGAGAGAGAGTGTCAAAAATATGACACAGTGTATCTTA 1347
QY 1081 CTTGTTCCAAAGGAAATCTCTGACAGAGACAAAGTCTTTGTTGAGTAATGCAAAATGCTGGT 1140
Db 1348 CTTGTTCCAAAGGAAATCTCTGACAGAGACAAAGTCTTTGTTGAGTAATGCAAAATGCTGGT 1407
QY 1141 GATTCGAAGCTGCGGAGAAATTAATGTAATGTAAGAGATGTTGTAATGGAAGTGTGTTG 1200
Db 1408 GATTCGAAGCTGCGGAGAAATTAATGTAATGTAAGAGATGTTGTAATGGAAGTGTGTTG 1467
QY 1201 GTACATTTACCCGCTGAGTGCCTACTCAAGAGGGGATATCTGTATCAAACTTGTCTTA 1260
Db 1468 GTACATTTACCCGCTGAGTGCCTACTCAAGAGGGGATATCTGTATCAAACTTGTCTTA 1527
RESULT 4
LOCUS HSA536197 2909 bp mRNA linear PRI 11-JAN-2003
DEFINITION Homo sapiens mRNA for KIS protein.
ACCESSION AJ536197
VERSION AJ536197.1 GI:27657360
KEYWORDS KIS protein.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Bieche, I., Manceau, V., Curmi, P. A., Laurendon, I., Lachkar, S.,
Leroy, K., Vadaud, D., Sobel, A. and Maucuer, A.
TITLE Quantitative RT-PCR reveals a ubiquitous but preferentially neural
expression of the KIS gene in rat and human
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 2909)
TITLE Direct Submision
JOURNAL Submitted (08-JAN-2003) Maucuer A., U440, Inserm, IFM, 17 rue du
fer a moulin, 75005 Paris, FRANCE
FEATURES
source
1. .2909
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/db_xref="taxon:9606"
/clone="I.M.A.G.E. Consortium CloneID 145548"
/issue_type="2 pooled tumors (clear cell type)"
1. .2909
/gene="KIS"
146. .1405
/gene="KIS"
/codon_start=1
/product="KIS protein"
/protein_id="CAD60192.1"
/db_xref="GI:27657361"
/translation="MAGSGCAGMBAPRFLAAGRLVYOSRLSGSSASVYRVCCG
NPGSPGALKQPLPFGTTGAASAAATGKRKRAALQJQGRNITLYLIGVTTIHSP
NPFSRLLELDVSVSELLVSHQSCMMMIQHCARVLEBALAHLEGVHADLK
PRLTWSANECEPKLIDFGLSPKBNQDVKYIQDQVRAPEALQCLAQAGVQSTE
CTSAVDLSLIGIILEMFSGMKLKHTVRSQEWKANSALIDHI PASKAVNNAIPAYH
LRLIKSMLHDDPSRRI PAMKLCSPEFSI PFAPIHEDIVMLPTVLRLLVLDLYL
ENEEEDYEDVEDYKESCKYKPVSLVLEKENGQGVVEYANAGDSRAQKLLTGR
MEDGRKVVATTFYLSAYKRGYLYQTL"

ORIGIN

Query Match 99.6%; Score 1255.2; DB 9; Length 2909;
Best Local Similarity 99.8%; Pred. No. 1.4e-310;
Matches 1257; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 146 ATGGCGGAGATCCGGCTGCGCCTG66GCGCGAGCCGCGCTTTTCTGAGGCGCTTCG66
QY 61 CGAGTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCTCGGCTGGTATCGGGT 120
Db 206 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCTCGGCTGGTATCGGGT 265
QY 121 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCTCGGCTGGTATCGGGT 180
Db 266 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCTCGGCTGGTATCGGGT 325
QY 181 ACCACCGGGGCTGGGCGCTCTGGGCGCGAGTAATGTTTCCGAAAGAGAGGGCGGCGT 240
Db 326 ACCACCGGGGCTGGGCGCTCTGGGCGCGAGTAATGTTTCCGAAAGAGAGGGCGGCGT 385
QY 241 GAACAGTTGACGGGTACAGAAATACGTGACTTTGTATGAGAGTGTACAACTCACTTT 300
Db 386 GAACAGTTGACGGGTACAGAAATACGTGACTTTGTATGAGAGTGTACAACTCACTTT 445
QY 301 TCTCCAAATGTGCATCAAGCTGTCTGTGCTTGAATCTTGGATGTCAATGTTTCGAA 360
Db 446 TCTCCAAATGTGCATCAAGCTGTCTGTGCTTGAATCTTGGATGTCAATGTTTCGAA 505
QY 361 TTGCTCTTATATTCAGATCACAGGCTTGTCCAGTGTGATGATACAGATTTGGCGCCGA 420
Db 506 TTGCTCTTATATTCAGATCACAGGCTTGTCCAGTGTGATGATACAGATTTGGCGCCGA 565
QY 421 GATGTTTGGAGGCGCTTGTCTTCTCATCATGAGGCTATGTGCTCATGCGGACTCAAA 480
Db 566 GATGTTTGGAGGCGCTTGTCTTCTCATCATGAGGCTATGTGCTCATGCGGACTCAAA 625
QY 481 CCACGTAAATATTTGTGAGTGCAGAGATGAAATGTTTAACTCATTTGACTTTGACTT 540
Db 626 CCACGTAAATATTTGTGAGTGCAGAGATGAAATGTTTAACTCATTTGACTTTGACTT 685
QY 541 ACCTTCAAGAAAGGCAATCAGATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 600
Db 686 ACCTTCAAGAAAGGCAATCAGATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 745
QY 601 GAAGCAGAAATTCGAAATTTGCTTGCGCAGGCTGCGAGAGTATAGAAATGTATAC 660
Db 746 GAAGCAGAAATTCGAAATTTGCTTGCGCAGGCTGCGAGAGTATAGAAATGTATAC 805
QY 661 TCAGCTGTTGATCTGTGAGGCTTGAAGATCAATTTTACTGAAATGTTCTCAGAAATGAA 720
Db 806 TCAGCTGTTGATCTGTGAGGCTTGAAGATCAATTTTACTGAAATGTTCTCAGAAATGAA 865
QY 721 CTGAACATACAGTACAGATCTCAGAAATGAAAGCAAGTTCGCTATTTATGATAC 780
Db 866 CTGAACATACAGTACAGATCTCAGAAATGAAAGCAAGTTCGCTATTTATGATAC 925
QY 781 ATATTGCGAGTAAGCAGTGGTGAATCGCAATTTCCAGCTATCACCTTAAGAGACCTT 840
Db 926 ATATTGCGAGTAAGCAGTGGTGAATCGCAATTTCCAGCTATCACCTTAAGAGACCTT 985
QY 841 ATCAAAGAGATGCTTCAATGATGATCAAGCAGAAAGATTTCTGCTGAATGGCATTTGTC 900
Db 986 ATCAAAGAGATGCTTCAATGATGATCAAGCAGAAAGATTTCTGCTGAATGGCATTTGTC 1045
QY 901 AGCCATTTCTTGAATCTTTTCCCTCATATTTGAAGATCTGTATGCTTCCACT 960
Db 1046 AGCCATTTCTTGAATCTTTTCCCTCATATTTGAAGATCTGTATGCTTCCACT 1105
QY 961 CCAGTGTCAAGCTGCTGAATGTGTGATGATTAATCTTGGAAATGAAGAGAAAT 1020
Db 1106 CCAGTGTCAAGCTGCTGAATGTGTGATGATTAATCTTGGAAATGAAGAGAAAT 1165
QY 1021 GAAGATGTTGTAGAAAGATGTAAGAGAGAGTGTCAAAAATATGACACAGTGTATCTTA 1080
Db 1166 GAAGATGTTGTAGAAAGATGTAAGAGAGAGTGTCAAAAATATGACACAGTGTATCTTA 1225
QY 1081 CTTGTTCCAAAGGAAATCTCTGACAGAGACAAAGTCTTTGTTGAGTATGCAATGCTGGT 1140
Db 1226 CTTGTTCCAAAGGAAATCTCTGACAGAGACAAAGTCTTTGTTGAGTATGCAATGCTGGT 1285

QY 1141 GATTCGAAGCTGCGAGAAATTAAGTGAAGAGATGTTGAGGAAAGTTGTTG 1200
DB 1286 GATTCGAAGCTGCGAGAAATTAAGTGAAGAGATGTTGAGGAAAGTTGTTG 1345
QY 1201 GCTACATTTACCCGCTGAGTGCCTACAGAGGGGATATCTGTATCAAACTTGTCTTAA 1260
DB 1346 GCTACATTTACCCGCTGAGTGCCTACAGAGGGGATATCTGTATCAAACTTGTCTTAA 1405

RESULT 5
CQ412057 4065 bp DNA linear PAT 23-JAN-2004
LOCUS Sequence 19128 from Patent WO0170979.
DEFINITION CQ412057
ACCESSION CQ412057
VERSION CQ412057.1 GI:41319838
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
AUTHORS Lee, J. and Lillie, J.
TITLE Genes, compositions, kits, and method for identification, assessment, prevention, and therapy of ovarian cancer
JOURNAL Patent: WO 0170979-A 19128 27-SEP-2001,
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source 1..4065
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.6%; Score 1255.2; DB 6; Length 4065;
Best Local Similarity 99.8%; Pred. No. 1.5e-310;
Matches 1257; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGCGCGGATCCGGCTGCGCTGGGGCGGAGCCCGCGTTTTCGAGGCGCTTCGAG 60
DB 268 ATGCGCGGATCCGGCTGCGCTGGGGCGGAGCCCGCGTTTTCGAGGCGCTTCGAG 327
QY 61 CGGCTGTCGAGGTACAGAGCCGCTGTCGTCGAGCGCTCTCCGCTCGGTATCGGGTT 120
DB 328 CGGCTGTCGAGGTACAGAGCCGCTGTCGTCGAGCGCTCTCCGCTCGGTATCGGGTT 387
QY 121 CGCTGTCGCGCAACCTGCTGCGCCCGCGGCGCCTTCAGGAGATTCTTGCGCGAGA 180
DB 388 CGCTGTCGCGCAACCTGCTGCGCCCGCGGCGCCTTCAGGAGATTCTTGCGCGAGA 447
QY 181 ACCACCGGGGCTGCGGCTTTCGCGGAGTATGTTTCCGCAAGAAGAGGGGCGGCTG 240
DB 448 ACCACCGGGGCTGCGGCTTTCGCGGAGTATGTTTCCGCAAGAAGAGGGGCGGCTG 507
QY 241 GACAGTTCGAGGGTTCACAGAAACATGTCATTTGATGAGGTTTACATCCACTTT 300
DB 508 GACAGTTCGAGGGTTCACAGAAACATGTCATTTGATGAGGTTTACATCCACTTT 567
QY 301 TCTCCAATGTGCATCAGCTGCTGTTGCTTGAACCTCTGATGTCAGTGTTCGAA 360
DB 568 TCTCCAATGTGCATCAGCTGCTGTTGCTTGAACCTCTGATGTCAGTGTTCGAA 627
QY 361 TTGCTCTTATATTCAGTCAACAGGGTTCCTCATGTGATGATACAGCATTCGCCCGA 420
DB 628 TTGCTCTTATATTCAGTCAACAGGGTTCCTCATGTGATGATACAGCATTCGCCCGA 687
QY 421 GATGTTTGAAGGCCCTTGTCTTTCTCATCATGAGGGCTATGTCATGCGGACCTCAA 480
DB 688 GATGTTTGAAGGCCCTTGTCTTTCTCATCATGAGGGCTATGTCATGCGGACCTCAA 747
QY 481 CCAAGTAACATATGTGAGTGCAGAGAAATGTTTAACTCATTTGAAGCTT 540
DB 748 CCAAGTAACATATGTGAGTGCAGAGAAATGTTTAACTCATTTGAAGCTT 807

QY 541 ACCTTCAAGAGGCAATCAGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
DB 808 ACCTTCAAGAGGCAATCAGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 867
QY 601 GAAGCAAAATTCAAAATTCCTTGGCCCAAGCTGGCTGCGAGTGTATACAGAAATGTAC 660
DB 868 GAAGCAAAATTCAAAATTCCTTGGCCCAAGCTGGCTGCGAGTGTATACAGAAATGTAC 927
QY 661 TCAGCTGTGATCTGTGAGGCTTAGAATCAATTTACTGAAATGTTCTCAGAAATGAAA 720
DB 928 TCAGCTGTGATCTGTGAGGCTTAGAATCAATTTACTGAAATGTTCTCAGAAATGAAA 987
QY 721 CTGAAACATACAGTCAAGTCTCAGAGATGAGAGCAACAGTTCTGTATTAATGATCAC 780
DB 988 CTGAAACATACAGTCAAGTCTCAGAGATGAGAGCAACAGTTCTGTATTAATGATCAC 1047
QY 781 ATATTTGCCAGTAAGCAAGTGTGAATGCCCAATTCAGGCTTATCAGCTTAAGAGACCTT 840
DB 1048 ATATTTGCCAGTAAGCAAGTGTGAATGCCCAATTCAGGCTTATCAGCTTAAGAGACCTT 1107
QY 841 ATCAAAAGATGCTTCATGATGATCAAGCAAGAAATTCCTGCTGAATGSCATTTGTC 900
DB 1108 ATCAAAAGATGCTTCATGATGATCAAGCAAGAAATTCCTGCTGAATGSCATTTGTC 1167
QY 901 AGCCCATTTCTTGAAGATTCCTTTGCCCCCTCATATTTGAAGATCTGTATGCTTCCACT 960
DB 1168 AGCCCATTTCTTGAAGATTCCTTTGCCCCCTCATATTTGAAGATCTGTATGCTTCCACT 1227
QY 961 CCAAGTCTAAGACTCTGTAATGCTGATGATGATATCTTGGGAATGAAGAGAAATAT 1020
DB 1228 CCAAGTCTAAGACTCTGTAATGCTGATGATGATATCTTGGGAATGAAGAGAAATAT 1287
QY 1021 GAAGATGTTGTGAAGATGTAAAGAGAGTCCAAAATATGACAGTGTATCTCTA 1080
DB 1288 GAAGATGTTGTGAAGATGTAAAGAGAGTCCAAAATATGACAGTGTATCTCTA 1347
QY 1081 CTTGTTCCAAAGAAATCTCTGCGAGAGCAAGCTTGTGTAATGCAAACTCTGGT 1140
DB 1348 CTTGTTCCAAAGAAATCTCTGCGAGAGCAAGCTTGTGTAATGCAAACTCTGGT 1407
QY 1141 GATTCGAAGCTGCGCAAAATTAAGTGAAGAGATGTTGATGGAAGTTTGTGTG 1200
DB 1408 GATTCGAAGCTGCGCAAAATTAAGTGAAGAGATGTTGATGGAAGTTTGTGTG 1467
QY 1201 GCTACATTTTACCCGCTGAGTGCCTTACAGAGGGGATATCTGTATCAAACTTGTCTTAA 1260
DB 1468 GCTACATTTTACCCGCTGAGTGCCTTACAGAGGGGATATCTGTATCAAACTTGTCTTAA 1527

RESULT 6
AX188296 4065 bp DNA linear PAT 06-AUG-2001
LOCUS AX188296
DEFINITION Sequence 3991 from Patent WO0142467.
ACCESSION AX188296
VERSION AX188296.1 GI:15139769
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
AUTHORS Schlegel, R., Deeds, J., Berger, A. and Zhao, X.
TITLE Genes, compositions, kits, and methods for identification, assessment, prevention, and therapy of cervical cancer
JOURNAL Patent: WO 0142467-A 3991 14-JUN-2001,
Millennium Predictive Medicine, Inc. (US)
FEATURES
source 1..4065
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 99.6%; Score 1255.2; DB 6; Length 4065;
 Best Local Similarity 99.8%; Pred. No. 1.5e-310;
 Matches 1257; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 ATGGCGGGAATCCGGCTCCGCTGGGGGCGCGAGACCCGCGCTTTTCTGAGAGGCTTCGGG 60
DB 268 ATGGCGGGAATCCGGCTCCGCTGGGGGCGCGAGACCCGCGCTTTTCTGAGAGGCTTCGGG 327
QY 61 CGGCTGTGAGGATACAGAGCCGCTCTGGGTAGCGGCTCTCCGCTCGGTGTATCGGGTT 120
DB 328 CGGCTGTGAGGATACAGAGCCGCTCTGGGTAGCGGCTCTCCGCTCGGTGTATCGGGTT 387
QY 121 CGGCTGTGCGGCAACCTGGCTGCCCCCGGCGCCCTCAGGAGTTCTTTCGCGCAGAA 180
DB 388 CGGCTGTGCGGCAACCTGGCTGCCCCCGGCGCCCTCAGGAGTTCTTTCGCGCAGAA 447
QY 181 ACCACCGGGGCTCGGCTCTGCGCCCGAGATAGTTTCCGCAAAAGAGAGGGCGGCTG 240
DB 448 ACCACCGGGGCTCGGCTCTGCGCCCGAGATAGTTTCCGCAAAAGAGAGGGCGGCTG 507
QY 241 GAACAGTTGAGGGGTCAAGAAAATGTGACTTTGTATGAGTGTATTAATCCACTT 300
DB 508 GAACAGTTGAGGGGTCAAGAAAATGTGACTTTGTATGAGTGTATTAATCCACTT 567
QY 301 TCTCCAAATGTGCATCAGCTGTCTGTGCTTGAACCTCTGGATGTCAAGTGTTCGAA 360
DB 568 TCTCCAAATGTGCATCAGCTGTCTGTGCTTGAACCTCTGGATGTCAAGTGTTCGAA 627
QY 361 TTGCTCTTATATTCAGATCAGAGGTTTTCATGTGATGTATACAGATTCGCGCCA 420
DB 628 TTGCTCTTATATTCAGATCAGAGGTTTTCATGTGATGTATACAGATTCGCGCCA 687
QY 421 GATGTTTGGAGGCGCTTGTCTTTCATCATAGAGGCTATGCTCATGCGAGCTCAAA 480
DB 688 GATGTTTGGAGGCGCTTGTCTTTCATCATAGAGGCTATGCTCATGCGAGCTCAAA 747
QY 481 CCAAGTAAATATTTGAGATGAGAGAAATGTTTAACTCATTTGACTT 540
DB 748 CCAAGTAAATATTTGAGATGAGAGAAATGTTTAACTCATTTGACTT 807
QY 541 AGCTTCAAGAGAGGCAATCAGATGTAAATATTTCAAGACAGGCTATCGGCTCCA 600
DB 808 AGCTTCAAGAGAGGCAATCAGATGTAAATATTTCAAGACAGGCTATCGGCTCCA 867
QY 601 GAAGCAGAAATTTGCAAAATTTGCTGGCCAGGCTGGCTGAGAGTATCAGAAATATAC 660
DB 868 GAAGCAGAAATTTGCAAAATTTGCTGGCCAGGCTGGCTGAGAGTATCAGAAATATAC 927
QY 661 TCAAGCTGTGATCTGTGAGCTTGAAGATCAATTTTACTGAAATGTTTCTCAGAAATGAA 720
DB 928 TCAAGCTGTGATCTGTGAGCTTGAAGATCAATTTTACTGAAATGTTTCTCAGAAATGAA 987
QY 721 CTGAAAATATCAGATCAGATCTCAGAAATGAGGCAACAGTTCTGTATTTATTTATCAC 780
DB 988 CTGAAAATATCAGATCAGATCTCAGAAATGAGGCAACAGTTCTGTATTTATTTATCAC 1047
QY 781 ATATTGCGCAGTAAGAGATGAGTGAATGCGCAATTCAGCCCTATACCTAAGAGACCTT 840
DB 1048 ATATTGCGCAGTAAGAGATGAGTGAATGCGCAATTCAGCCCTATACCTAAGAGACCTT 1107
QY 841 ATCAAAAGCATGCTTATGATGATTCAGAGCAGAAAGATTCCTGCTGAAATGAGATTTGTC 900
DB 1108 ATCAAAAGCATGCTTATGATGATTCAGAGCAGAAAGATTCCTGCTGAAATGAGATTTGTC 1167
QY 901 AGCCCAATTTTATGATCTTTTTCGCCCTCATATTGAAGATGTGCTATGCTCCACT 960
DB 1168 AGCCCAATTTTATGATCTTTTTCGCCCTCATATTGAAGATGTGCTATGCTCCACT 1227
QY 961 CCAAGTCAAGATGCTGAGATGCTGAGATGATGATTAATCTGGGAAATGAAGAGAAAT 1020
DB 1228 CCAAGTCAAGATGCTGAGATGCTGAGATGATGATTAATCTGGGAAATGAAGAGAAAT 1287

```

```

QY 1021 GAAGATGTTGAGAGATGTAAAGAGAGTGTCAAAAATATGACCAAGTGTATCTCTA 1080
DB 1288 GAAGATGTTGAGAGATGTAAAGAGAGTGTCAAAAATATGACCAAGTGTATCTCTA 1347
QY 1081 CTTGTTCCAAAGAAATCTTGGCGAGAGACAAAGCTTTGTTAGATAGCAAAATGCTGT 1140
DB 1348 CTTGTTCCAAAGAAATCTTGGCGAGAGACAAAGCTTTGTTAGATAGCAAAATGCTGT 1407
QY 1141 GATTCCAAAGCTGCGCAGAAATTAATTAATGACTGAGAGATGTTTATGAGAAATGTTGTG 1200
DB 1408 GATTCCAAAGCTGCGCAGAAATTAATTAATGACTGAGAGATGTTTATGAGAAATGTTGTG 1467
QY 1201 GCTAATTTATCCCGCTAGTGTCTTACAAAGAGGAGATATCTGTATCAAACTTGTCTTAA 1260
DB 1468 GCTAATTTATCCCGCTAGTGTCTTACAAAGAGGAGATATCTGTATCAAACTTGTCTTAA 1527

```

RESULT 7
 AX16543 1260 bp DNA linear PAT 22-JUN-2001
 LOCUS AX16543
 DEFINITION Sequence 34 from Patent WO0138503.
 ACCESSION AX16543
 VERSION AX16543.1 GI:14546888
 KEYWORDS

SOURCE
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 Plozman, G.D., Whyte, D., Manning, G.S., Sudarsanam, S.S., Martinez, R.,
 Flanagan, P., and Clardy, D.S.
 TITLE Novel human protein kinases and protein kinase-1-like enzymes
 JOURNAL Patent: WO 0138503-A 34 31-MAY-2001;
 Sugan, Inc. (US)

FEATURES
 source location/Qualifiers
 1..1260
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

Query Match 99.5%; Score 1253.6; DB 6; Length 1260;
 Best Local Similarity 99.7%; Pred. No. 3e-310;
 Matches 1256; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 ATGGCGGGAATCCGGCTCCGCTGGGGGCGCGAGACCCGCGCTTTTCTGAGAGGCTTCGGG 60
DB 1 ATGGCGGGAATCCGGCTCCGCTGGGGGCGCGAGACCCGCGCTTTTCTGAGAGGCTTCGGG 60
QY 61 CGGCTGTGAGGATACAGAGCCGCTCTGGGTAGCGGCTCTCCGCTCGGTGTATCGGGTT 120
DB 61 CGGCTGTGAGGATACAGAGCCGCTCTGGGTAGCGGCTCTCTCCGCTCGGTGTATCGGGTT 120
QY 121 CGGCTGTGCGGCAACCTGGCTGCCCCCGGCGCCCTCAGGAGTTCTTTCGCGCAGAA 180
DB 121 CGGCTGTGCGGCAACCTGGCTGCCCCCGGCGCCCTCAGGAGTTCTTTCGCGCAGAA 180
QY 181 ACCACCGGGGCTCGGCTCTGCGCCCGAGATAGTTTCCGCAAAAGAGAGGGCGGCTG 240
DB 181 ACCACCGGGGCTCGGCTCTGCGCCCGAGATAGTTTCCGCAAAAGAGAGGGCGGCTG 240
QY 241 GAACAGTTGAGGGGTCAAGAAAATGTGACTTTGTATGAGTGTATTAATCCACTT 300
DB 241 GAACAGTTGAGGGGTCAAGAAAATGTGACTTTGTATGAGTGTATTAATCCACTT 300
QY 301 TCTCCAAATGTGCATCAGCTGTCTGTGCTTGAACCTCTGGATGTCAAGTGTTCGAA 360
DB 301 TCTCCAAATGTGCATCAGCTGTCTGTGCTTGAACCTCTGGATGTCAAGTGTTCGAA 360
QY 361 TTGCTCTTATATTCAGATCAGAGGTTTTCATGTGATGTATACAGATTCGCGCCA 420
DB 361 TTGCTCTTATATTCAGATCAGAGGTTTTCATGTGATGTATACAGATTCGCGCCA 420

```



```

Oy 421 GATGTTTGGAGGCCCTTCTTCTTTCATCATGAGGGCTATGTCATGCGGACCTCAAA 480
Db 421 GATGTTTGGAGGCCCTTCTTCTTTCATCATGAGGGCTATGTCATGCGGACCTCAAA 480
Oy 481 CCACGTACATATGTTGAGTGCAGAGAAATGAAATGTTTAACTCATGACTTTGGACTT 540
Db 481 CCACGTACATATGTTGAGTGCAGAGAAATGAAATGTTTAACTCATGACTTTGGACTT 540
Oy 541 ACCTTCAAGAGGAGCATGAGATGTAAGTATATTCAGACAGACGGGATGCGGCTCCA 600
Db 541 ACCTTCAAGAGGAGCATGAGATGTAAGTATATTCAGACAGACGGGATGCGGCTCCA 600
Oy 601 GAAGCAGAAATTCGAAATTCCTTGGCCCAAGCTGCGCTCAGAGTATACAGATATACC 660
Db 601 GAAGCAGAAATTCGAAATTCCTTGGCCCAAGCTGCGCTCAGAGTATACAGATATACC 660
Oy 661 TCAGCTGTTGATCTGTGAGCTGAGGATCATTTTAACTGAAATGTTCTCAGGAATGAAA 720
Db 661 TCAGCTGTTGATCTGTGAGCTGAGGATCATTTTAACTGAAATGTTCTCAGGAATGAAA 720
Oy 721 CTGAAAACATACATGATCTGAGGAAATGGAAGGCAACAGTTCTGCTATTTATGATCAC 780
Db 721 CTGAAAACATACATGATCTGAGGAAATGGAAGGCAACAGTTCTGCTATTTATGATCAC 780
Oy 781 ATATTGCGCAGTAAAGCAGTGTGAAATGCGCAATTCAGACCTATACCTTAAGAGACTT 840
Db 781 ATATTGCGCAGTAAAGCAGTGTGAAATGCGCAATTCAGACCTATACCTTAAGAGACTT 840
Oy 841 ATCAAAAGCATGCTTCATGATGATCCAGACAGAAAGATTCCTGCTGAATGCGACTTGTG 900
Db 841 ATCAAAAGCATGCTTCATGATGATCCAGACAGAAAGATTCCTGCTGAATGCGACTTGTG 900
Oy 901 AGCCCATTTCTTAAAGCATTCCTTTTGGCCCTCATATGGAAGATCTGCTCATCTCCACT 960
Db 901 AGCCCATTTCTTAAAGCATTCCTTTTGGCCCTCATATGGAAGATCTGCTCATCTCCACT 960
Oy 961 CCAGTCTAAGACTGCTGATGATGCTGATGATGATTTATCTTGGGAATGAAGAGAAATAT 1020
Db 961 CCAGTCTAAGACTGCTGATGATGCTGATGATGATTTATCTTGGGAATGAAGAGAAATAT 1020
Oy 1021 GAAGATGTTGTAGAGATGTAAAGAGAGAGATCAAAAATATGAGACAGTGTATCTCTA 1080
Db 1021 GAAGATGTTGTAGAGATGTAAAGAGAGAGATCAAAAATATGAGACAGTGTATCTCTA 1080
Oy 1081 CTTGTTTCCAAAGGAAATCTCTGCGACAGACAGATCTTGTGAGATATCAAAATGCTGCT 1140
Db 1081 CTTGTTTCCAAAGGAAATCTCTGCGACAGACAGATCTTGTGAGATATCAAAATGCTGCT 1140
Oy 1141 GATTCGAAGCTGCGAGAAATTAATGACTGAGAGAGATGTTGATGGAAGTTGTTGTG 1200
Db 1141 GATTCGAAGCTGCGAGAAATTAATGACTGAGAGAGATGTTGATGGAAGTTGTTGTG 1200
Oy 1201 GCTACATTTACCCGCTGAGTGCCTAACAAGGGGATATCTGTATCAAACTTGTCTTAA 1260
Db 1201 GCTACATTTACCCGCTGAGTGCCTAACAAGGGGATATCTGTATCAAACTTGTCTTAA 1260

```

```

RESULT 8
AX680136 1824 bp DNA linear PAT 29-MAR-2003
LOCUS Sequence 21 from Patent WO0181555.
DEFINITION AX680136
ACCESSION AX680136
VERSION AX680136.1 GI:29369916
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE
AUTHORS
1 Gururajan,R., Paterson,C., Recipon,S.A., Tribouley,C.M.,
Hafalla,A., Khan,F.Y., Yue,H., Au-Young,J., Bandman,O.,
Baughn,M.R., Borowsky,M.L., Buford,N., Burdill,J.D., Elliott,V.S.,
Gandhi,A.R., Kearney,L., Lal,P., Lu,D.A., Lu,Y., Tang,Y.T.,

```

```

Zingler,K.A., Griffin,J.A., Hillman,J.J., Marcus,G.A., Nguyen,D.B.,
Policky,J., Ramkumar,J., Thornton,M., Wallis,N.K. and Walsh,R.T.
Human Kinases
Patent: NO 0181555-A 21 01-NOV-2001;
Incyte Genomics, Inc. (US)
FEATURES
source
1..1824
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 3107952CB1"
ORIGIN
Query Match 99.5%; Score 1253.6; DB 6; Length 1824;
Best Local Similarity 99.7%; Pred. No. 3.3e-310;
Matches 1256; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

Oy 1 ATGGCGGAGATCCGCTGCGCTGCGGAGCGGAGCGGCGGCTTCTGAGGCTTGGG 60
Db 1 ATGGCGGAGATCCGCTGCGCTGCGGAGCGGAGCGGCGGCTTCTGAGGCTTGGG 60
Oy 169 ATGGCGGAGATCCGCTGCGCTGCGGAGCGGAGCGGCGGCTTCTGAGGCTTGGG 228
Db 169 ATGGCGGAGATCCGCTGCGCTGCGGAGCGGAGCGGCGGCTTCTGAGGCTTGGG 228
Oy 61 CGGCTGTGCGAGGTACAGAGCCTGTGGGTAGCGGCTCTCGCTCGGTGTATCGGTT 120
Db 61 CGGCTGTGCGAGGTACAGAGCCTGTGGGTAGCGGCTCTCGCTCGGTGTATCGGTT 120
Oy 229 CGGCTGTGCGAGGTACAGAGCCTGTGGGTAGCGGCTCTCGCTCGGTGTATCGGTT 288
Db 229 CGGCTGTGCGAGGTACAGAGCCTGTGGGTAGCGGCTCTCGCTCGGTGTATCGGTT 288
Oy 121 CGCTGCTGGGCAACCTGAGCTGCGCCCGGAGGCTCAGGAGATTCTTGCGGCAAGA 180
Db 121 CGCTGCTGGGCAACCTGAGCTGCGCCCGGAGGCTCAGGAGATTCTTGCGGCAAGA 180
Oy 288 CGCTGCTGGGCAACCTGAGCTGCGCCCGGAGGCTCAGGAGATTCTTGCGGCAAGA 348
Db 288 CGCTGCTGGGCAACCTGAGCTGCGCCCGGAGGCTCAGGAGATTCTTGCGGCAAGA 348
Oy 181 ACCACCGGGGCTGCGGCTCTGCGCGCAGATATGTTTCCGCAAGAGAGGGCGGCTG 240
Db 181 ACCACCGGGGCTGCGGCTCTGCGCGCAGATATGTTTCCGCAAGAGAGGGCGGCTG 240
Oy 349 ACCACCGGGGCTGCGGCTCTGCGCGCAGATATGTTTCCGCAAGAGAGGGCGGCTG 408
Db 349 ACCACCGGGGCTGCGGCTCTGCGCGCAGATATGTTTCCGCAAGAGAGGGCGGCTG 408
Oy 241 GAACGTTGCGAGGTACAGAAACATGCTGACTTTGATGAGATTTTACATCCACTTT 300
Db 241 GAACGTTGCGAGGTACAGAAACATGCTGACTTTGATGAGATTTTACATCCACTTT 300
Oy 409 GAACGTTGCGAGGTACAGAAACATGCTGACTTTGATGAGATTTTACATCCACTTT 468
Db 409 GAACGTTGCGAGGTACAGAAACATGCTGACTTTGATGAGATTTTACATCCACTTT 468
Oy 301 TCTCAAAATGTCATCAAGCTGCTGTTGCTGAAACCTGGAATGCAATGTTTGGAA 360
Db 301 TCTCAAAATGTCATCAAGCTGCTGTTGCTGAAACCTGGAATGCAATGTTTGGAA 360
Oy 469 TCTCAAAATGTCATCAAGCTGCTGTTGCTGAAACCTGGAATGCAATGTTTGGAA 528
Db 469 TCTCAAAATGTCATCAAGCTGCTGTTGCTGAAACCTGGAATGCAATGTTTGGAA 528
Oy 361 TTGCTCTTATATTCAGTACACAGAGGTTTTCATGATGATATACAGACTTGCAGCA 420
Db 361 TTGCTCTTATATTCAGTACACAGAGGTTTTCATGATGATATACAGACTTGCAGCA 420
Oy 529 TTGCTCTTATATTCAGTACACAGAGGTTTTCATGATGATATACAGACTTGCAGCA 588
Db 529 TTGCTCTTATATTCAGTACACAGAGGTTTTCATGATGATATACAGACTTGCAGCA 588
Oy 421 GATGTTTGGAGGCCCTTCTTCTTTCATCATGAGGGCTATGTCATGCGGACCTCAAA 480
Db 421 GATGTTTGGAGGCCCTTCTTCTTTCATCATGAGGGCTATGTCATGCGGACCTCAAA 480
Oy 589 GATGTTTGGAGGCCCTTCTTCTTTCATCATGAGGGCTATGTCATGCGGACCTCAAA 648
Db 589 GATGTTTGGAGGCCCTTCTTCTTTCATCATGAGGGCTATGTCATGCGGACCTCAAA 648
Oy 481 CCACGTACATATGTTGAGTGCAGAGAAATGTTTAACTGATGACTTGGACTT 540
Db 481 CCACGTACATATGTTGAGTGCAGAGAAATGTTTAACTGATGACTTGGACTT 540
Oy 649 CCACGTACATATGTTGAGTGCAGAGAAATGTTTAACTGATGACTTGGACTT 708
Db 649 CCACGTACATATGTTGAGTGCAGAGAAATGTTTAACTGATGACTTGGACTT 708
Oy 541 ACCTTCAAGAGGAGCATCAGATGTAAATATTCAGACAGACGGGATGCGGCTCCA 600
Db 541 ACCTTCAAGAGGAGCATCAGATGTAAATATTCAGACAGACGGGATGCGGCTCCA 600
Oy 709 ACCTTCAAGAGGAGCATCAGATGTAAATATTCAGACAGACGGGATGCGGCTCCA 768
Db 709 ACCTTCAAGAGGAGCATCAGATGTAAATATTCAGACAGACGGGATGCGGCTCCA 768
Oy 601 GAAGCAGAAATTCGAAATTCCTTGGCCCAAGCTGCGCTCAGAGTATACAGATATACC 660
Db 601 GAAGCAGAAATTCGAAATTCCTTGGCCCAAGCTGCGCTCAGAGTATACAGATATACC 660
Oy 769 GAAGCAGAAATTCGAAATTCCTTGGCCCAAGCTGCGCTCAGAGTATACAGATATACC 828
Db 769 GAAGCAGAAATTCGAAATTCCTTGGCCCAAGCTGCGCTCAGAGTATACAGATATACC 828
Oy 661 TCAGCTGTTGATCTGTGAGCTGAGGATCATTTTAACTGAAATGTTCTCAGGAATGAAA 720
Db 661 TCAGCTGTTGATCTGTGAGCTGAGGATCATTTTAACTGAAATGTTCTCAGGAATGAAA 720
Oy 829 TCAGCTGTTGATCTGTGAGCTGAGGATCATTTTAACTGAAATGTTCTCAGGAATGAAA 888
Db 829 TCAGCTGTTGATCTGTGAGCTGAGGATCATTTTAACTGAAATGTTCTCAGGAATGAAA 888
Oy 721 CTGAAAACATACATGATCTGAGGAAATGGAAGGCAACAGTTCTGCTATTTATGATCAC 780
Db 721 CTGAAAACATACATGATCTGAGGAAATGGAAGGCAACAGTTCTGCTATTTATGATCAC 780
Oy 889 CTGAAAACATACATGATCTGAGGAAATGGAAGGCAACAGTTCTGCTATTTATGATCAC 948
Db 889 CTGAAAACATACATGATCTGAGGAAATGGAAGGCAACAGTTCTGCTATTTATGATCAC 948
Oy 781 ATATTGCGCAGTAAAGCAGTGTGAAATGCGCAATTCAGACCTATACCTTAAGAGACTT 840
Db 781 ATATTGCGCAGTAAAGCAGTGTGAAATGCGCAATTCAGACCTATACCTTAAGAGACTT 840
Oy 949 ATATTGCGCAGTAAAGCAGTGTGAAATGCGCAATTCAGACCTATACCTTAAGAGACTT 1008
Db 949 ATATTGCGCAGTAAAGCAGTGTGAAATGCGCAATTCAGACCTATACCTTAAGAGACTT 1008

```

Qy		841	ATCGAAGCATGCTTCATGTGAATGCCAGACAGAAGAAATCCGTGGTAATGGCANTTGGC	900
Db		1009	ATCAAAAAGCAAGCTTCATGATGATCCCAAGACAAGAATAATTCTGCTGAATATGCACATTGGC	1068
Qy		901	AGCCCATTTCTTAGCATCTCTTTTGCCCCCTCATATTGAAGATCTGTGATGCTTCCACT	960
Db		1069	AGCCCATTCTTTAGCATTTCTTTTTGCCCTCATNTGANAGNCTGGTCAATGCTTCCACT	1128
Qy		961	CCAAGTCTAAGACTGTCTGAATATGTGCTGGATGATGATTAATCTTTGGAAATGAAGAGATAT	1020
Db		1129	CAAAGTCTAAGACTGTCTGAATATGTGCTGGATGATGATTAATCTTTGGAAATGAAGAGATAT	1188
Qy		1021	GAAATGTTGNAGAAAGATNGTAAAAAGAGAGTGTCAAAAAATATGACCAGTGTATCTCTA	1080
Db		1189	GAAATGTTGNAGAAAGATNGTAAAAAGAGAGTGTCAAAAAATATGACCAGTGTATCTCTA	1248
Qy		1081	CTTGTTCGAAAGAAAATCTCTGGCAGAGACAAGTCTTTGTTGATATGCAATAGCTGGT	1140
Db		1249	CTTGTTCGAAAGGAAAATCTCTGGCAGAGACAAGTCTTTGTTGATATGCAAAATGCTGGT	1308
Qy		1141	GATTCCAAAGCTGCGCAGAAAATTACTGACTGGAAGAGATGTTTGATGGAACTTTGTGTG	1200
Db		1309	GATTCCAAAGCTGCGCAGAAAATTACTGACTGGAAGAGATGTTTGATGGAACTTTGTGTG	1368
Qy		1201	GCTACATTTCTTACCCGCTGAGTGCCCTAACAGAGGGGATATCTGTATCAAACTTGCTTAA	1260
Db		1369	GCTACATTTCTTACCCGCTGAGTGCCCTAACAGAGGGGATATCTGTATCAAACTTGCTTAA	1428
RESULT_9				
LOCUS	BC058732			
DEFINITION	Mus musculus kinase interacting with leukemia-associated gene (seafemin), mRNA (CDNA clone MGC:64770 IMAGE:6414877), complete cds.			
VERSION	BC058732			
KEYWORDS	BC058732.1 GI:37194892			
SOURCE	Mus musculus			
ORGANISM	Mus musculus (house mouse)			
REFERENCE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 2041)			
AUTHORS	Straubberg,R.L., Felingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shemen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Butow,K.H., Schafer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marsina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stetson,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,W.J., Uedin,T.B., Tothyluki,S., Canciani,P., Prange,C., Rana,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Wolley,K.C., Hale,S., Garcia,A.M., Gay,L.U., Hilyk,S.W., Villalón,D.K., Kozny,D.M., Sodergren,E.O., Lu,X., Gibbs,R.A., Fahney,J., Helton,E., Keltman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchan,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimmelwood,J., Schnurdt,J., Myers,R.M., Butterfield,Y.S., Krzyzaniecki,M.I., Skalska,U., Smilins,D.E., Schmerchig,A., Schein,J.E., Jones,S.J., and Matra,M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) 12477932			
TITLE	(bases 1 to 2041)			
JOURNAL	Strauberg,R.			
PUBMED	Submitted (24-Sep-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,			
REFERENCES	USA NIH-MGC Project URL: http://mgc.nci.nih.gov			
AUTHORS				
TITLE				
JOURNAL				
REMARK				

COMMENT

Contact: MGC help desk
 Email: gcaps-r@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: M. Bento Soares, University of Iowa
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant.
 Web site: <http://genome.uiowa.edu>
 Contact: bentos@uiowa.edu, tom-casavant@uiowa.edu
 Bonaldo, M.F., Akabogu, I., Balt, T., Balt, J., Crouch, K., Davis, A.,
 Flehler, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K.,
 Scheetz, T., Smith, C., Snir, E., Tack, D., Trout, K., Walters, J.,
 Casavant, T., Soares, M.B.

FEATURES

source

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: Plate: Row: Column: 0
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6754447.
 Location/Qualifiers

1..2041

/organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="MGC:64770 IMAGE:6414877"
 /tissue_type="Brain, pool of enriched mouse brain tissues from 13.5, 14.5, 16.5 and 17.5 dpc"
 /clone_lib="NIH BMAP_FRO"
 /lab_host="DH10B"
 /note="Vector: pYX-ASC"

1..2041

/gene="Kist"
 /note="synonym: KIS"
 /db_xref="LocusID:16589"
 /db_xref="MGI:1341908"
 154..1413

/gene="Kist"
 /codon_start=1
 /product="Kinase interacting with leukemia-associated gene (srcatmin)"
 /protein_id="AA58732.1"
 /db_xref="MGI:1341908"
 /db_xref="LocusID:16589"
 /db_xref="MGI:1341908"
 /translation="MAGSGCGAAGAEPRFLFLEAFRLMVOVSRIGSGSSAVRYRCCG
 TPSPGALAKQFLPGGTGTAAGAAGAEPRFKEPALEOLQGRNITVLYGVFTIHESP
 NVSPSCILFLDVSVELLYSSHSCSMNIOHCAVDLELAFLHGGVYADLK
 PRNIIWAENECFKLLIDFGISPRKGNQDVYIOTDGRAPAEIQLAAGLOSTPE
 CTAADVDSLGIILLLEFSGMKIKHTYRSGEMKANSALITDHTPASKAYNALIYAH
 LRLILSMILHDDPGRRIPAEMALCSPPFSIPFAPHLEMLTTPVLRILNLTDDYLL
 ENDEDEVDVDEKCEQKYPVSLVLPENPGRGVFEVYANAGDSKAAQKLLTER
 MFDGKVFATFVFLSAVKRGYLYQTL"

ORIGIN

Query Match 88.4%; Score 1114.4; DB 10; Length 2041;
 Best Local Similarity 92.8%; Pred. No. 1,7e-274;
 Matches 1169; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

Dh 1 ATGGCGGAGATCCGCTGCGCTGGCGGCGCGAGCGCGCGGTTTCTGAGAGCCTTGGG 60
 |||||
 154 ATGGCGGAGATCCGCTGCGCTGGCGGCGCGAGCGCGCGCTTCTGAGAGCCTTGGG 213
 |||||

Dh 61 CGGCTGTGAGAGGTACAGAGCCCTCTGTGGGTAGCGGCTCTCCGCTCGGTGTATCGGATT 120
 |||||
 214 CGGCTGTGAGAGGTACAGAGCCCTGTGGGCGAGCGGCTCGTCTCGGTGTATCGGATTG 273
 |||||

Dh 121 CGGCTGTGCGGCAACCTTGCGTGGCCCCCGCGCGCCCTCAGGAGAGTTCTTGGCCCGGAGA 180
 |||||
 274 CGGTGTGCGGTATCCCGGGGCTCGCCCCCGGGCGCCCTCAGGAGAGTTCTTGGCCCGGGA 333
 |||||

Dh 181 ACACCGGGGCTGGGGGCTCTGCGCGCGAGATAGTTTCCGCAAGAGAGGGCGGGCTG 240
 |||||

Db 334 ACCACCGGGGCTGGGCTTGGCCGGAGTATGTTTCGCAAGAGGGGGGCTG 393
 QY 241 GAACAGTTGAGGGGTCAAGAAACATGTCATCTTGTATGAGTGTATTAACATCACTTT 300
 Db 394 GAGCAATGAGGGGTCAAGAAACATGTCATCTTGTATGAGTGTATTAACATCACTTT 453
 QY 301 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 360
 Db 454 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 513
 QY 361 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 420
 Db 514 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 573
 QY 421 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 480
 Db 574 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 633
 QY 481 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 540
 Db 634 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 693
 QY 541 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 600
 Db 694 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 753
 QY 601 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 660
 Db 754 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 813
 QY 661 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 720
 Db 814 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 873
 QY 721 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 780
 Db 874 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 933
 QY 781 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 840
 Db 934 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 993
 QY 841 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 900
 Db 994 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 1053
 QY 901 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 960
 Db 1054 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 1113
 QY 961 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 1020
 Db 1114 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 1173
 QY 1021 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 1080
 Db 1174 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 1233
 QY 1081 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 1140
 Db 1234 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 1293
 QY 1141 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 1200
 Db 1294 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 1353
 QY 1201 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 1260
 Db 1354 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 1413

RESULT 10
 AC129141_1/c

WPCOMMENT
 Sequence split into 4 fragments LOCUS AC129141 Accession AC129141
 Fragment Name Begin End
 AC129141_0 1 110000
 AC129141_1 100001 210000
 AC129141_2 200001 310000
 AC129141_3 300001 363206
 Continuation (2 of 4) of AC129141 from base 100001 (AC129141 Rattus norvegicus clone CH2)

Query Match 88.2%; Score 1111.2; DB 2; Length 110000;
 Best Local Similarity 92.6%; Pred. No. 2.5e-273;
 Matches 1167; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 1 ATGGCGGAGATCCGGCTGCGCTGCGGCGGAGCGCGGCTTTCTGGAGGCTTTCGGG 60
 Db 70316 ATGGCGGAGATCCGGCTGCGCTGCGGCGGAGCGCGGCTTTCTGGAGGCTTTCGGG 70257
 QY 61 CGGCTGTGGAGGTACAGAGCCGTGTGAGTACGGCTCTCCGCTCGGTGTATCGGTT 120
 Db 70256 CGGCTGTGGAGGTACAGAGCCGTGTGAGTACGGCTCTCCGCTCGGTGTATCGGTT 70197
 QY 121 CGGCTGTGGAGGTACAGAGCCGTGTGAGTACGGCTCTCCGCTCGGTGTATCGGTT 180
 Db 70196 CGGCTGTGGAGGTACAGAGCCGTGTGAGTACGGCTCTCCGCTCGGTGTATCGGTT 70137
 QY 181 ACCACCGGGGCTGGCGGCTCTGCGCGGAGTATGTTTCCGCAAGAGAGGGCGGCTG 240
 Db 70136 ACCACCGGGGCTGGCGGCTCTGCGCGGAGTATGTTTCCGCAAGAGAGGGCGGCTG 70077
 QY 241 GAACAGTTGAGGGGTCAAGAAACATGTCATCTTGTATGAGTGTATTAACATCACTTT 300
 Db 70076 GAACAGTTGAGGGGTCAAGAAACATGTCATCTTGTATGAGTGTATTAACATCACTTT 70017
 QY 301 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 360
 Db 70016 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 69957
 QY 361 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 420
 Db 69956 TCTCAATGAGGGGTCAAGTGTATGAGTGTATTAACATCACTTT 69897
 QY 421 GATGTTTGGAGGCTTCTCTTTCTTCAATCAAGAGGCTATGTCATGCGGACTCAAA 480
 Db 69896 GATGTTTGGAGGCTTCTCTTTCTTCAATCAAGAGGCTATGTCATGCGGACTCAAA 69837
 QY 481 CCACTTAACATGTTGTGAGTGTCAAGAAATGATTTTAACTATGACTTTGACTT 540
 Db 69836 CCACTTAACATGTTGTGAGTGTCAAGAAATGATTTTAACTATGACTTTGACTT 69777
 QY 541 ACCTTCAAGAGGCAATCAGATGTTAAAGTATTTCAAGAGAGGGGATGCGGCTCCA 600
 Db 69776 ACCTTCAAGAGGCAATCAGATGTTAAAGTATTTCAAGAGAGGGGATGCGGCTCCT 69717
 QY 601 GAAGCAATGTCAAATGCTTGGCCGAGGCTGCGTCAAGAGTATCAAGAAATGTAAC 660
 Db 69716 GAAGCAATGTCAAATGCTTGGCCGAGGCTGCGTCAAGAGTATCAAGAAATGTAAC 69657
 QY 661 TCAAGCTGTGATCTGTGAGCTTGAAGTATCTTGAATCTGGAATGTTCTCAGAAATGAAA 720
 Db 69656 TCAAGCTGTGATCTGTGAGCTTGAAGTATCTTGAATCTGGAATGTTCTCAGAAATGAAA 69597
 QY 721 CTGAAACATGAGTCAAGTATCTCAGAAATGAGAAACAGTTCTGCTATTAATGATAC 780
 Db 69596 CTGAAACATGAGTCAAGTATCTCAGAAATGAGAAACAGTTCTGCTATTAATGATAC 69537
 QY 781 ATATTGCGAGTAAAGCAATGAGTGTGATCGCGCAATTTCAAGCTTATCACTTAAGACCTT 840
 Db 69536 ATATTGCGAGTAAAGCAATGAGTGTGATCGCGCAATTTCAAGCTTATCACTTAAGACCTT 69477
 QY 841 ATCAAAAGCATCTTCATGATGATCAAGCAAGAAATTTCTGCTGAATGCAATTTGTC 900
 Db 69476 ATCAAAAGCATCTTCATGATGATCAAGCAAGAAATTTCTGCTGAATGCAATTTGTC 69417

Db 999 CCAGTCTCACTCTCTCAATGTCTGATGATGATCTTGAAGAAATGATGATAT 1058
 QY 1021 GAAGATGTTGTAGAGATGTAAAGAGAGTCAAAAATATGACAGTGTATCTTA 1080
 Db 1059 GAAGATGTTGTAGAGATGTAAAGAGAGTCAAAAATATGACAGTGTATCTTA 1118
 QY 1081 CTGTGTTCAAGAGAAATCTGTGACAGAGCAAGTCTTTGTTGATGATGAAATGCTGT 1140
 Db 1119 CTGTGTTCAAGAGAAATCTGTGACAGAGCAAGTCTTTGTTGATGATGAAATGCTGT 1178
 QY 1141 GATTCAGAGCTGCGAGAGAAATCTGATCTGAGAGATGTTTATGAGAGAGTTGTTG 1200
 Db 1179 GATTCAGAGCTGCGAGAGAAATCTGATCTGAGAGATGTTTATGAGAGAGTTGTTG 1238
 QY 1201 GCTACATTCACCCGCTGAGTGTCTTCAAGAGAGGATATCTGATCAAACTTCTTAA 1260
 Db 1239 GCTACATTCACCCGCTGAGTGTCTTCAAGAGAGGATATCTTATCAAACTTCTTAA 1298
 RESULT 12 3244 bp mRNA linear ROD 08-SEP-1999
 LOCUS RN70372
 DEFINITION Rattus norvegicus PAM COOH-terminal interactor protein 2 mRNA,
 complete cds.
 ACCESSION U70372.2 GI:5821767
 VERSION U70372
 KEYWORDS
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 3244)
 AUTHORS Alam,M.R., Caldwell,B.D., Johnson,R.C., Darlington,D.N., Mains,R.E.
 TITLE Novel proteins that interact with the COOH-terminal cytosolic
 routing determinants of an integral membrane peptide-processing
 enzyme
 JOURNAL J. Biol. Chem. 271 (45), 28636-28640 (1996)
 MEDLINE 97067094
 PUBMED 8910496
 REFERENCE 2 (bases 1 to 3244)
 AUTHORS Caldwell,B.D., Darlington,D.N., Penzes,P., Johnson,R.C.,
 Bipper,B.A. and Mains,R.E.
 TITLE The Novel Kinase P-CIP2 Interacts with the Cytosolic Routing
 Determinants of the Peptide Processing Enzyme Peptidylglycine
 alpha-Amidating Monooxygenase
 JOURNAL J. Biol. Chem. 274 (1999) In press
 REFERENCE 3 (bases 1 to 3244)
 AUTHORS Alam,R., Caldwell,B.D., Johnson,R.C., Darlington,D.N., Mains,R.E.
 TITLE Direct Submission
 JOURNAL Submitted (11-SEP-1996) Neuroscience, Johns Hopkins University
 School of Medicine, 725 N. Wolfe Street, WBSB 902A, Baltimore, MD
 21205, USA
 REFERENCE 4 (bases 1 to 3244)
 AUTHORS Caldwell,B.D., Darlington,D.N., Penzes,P., Johnson,R.C.,
 Bipper,B.A. and Mains,R.E.
 TITLE Direct Submission
 JOURNAL Submitted (03-SEP-1999) Neuroscience, Johns Hopkins University
 School of Medicine, 725 N. Wolfe Street, WBSB 902A, Baltimore, MD
 21205, USA
 REMARK Sequence update by submitter
 COMMENT On Sep 3, 1999 this sequence version replaced gi:1698778.
 FEATURES
 source
 1..3244
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 162..11421
 /function="serine protein kinase"
 /note="P-CIP2"
 /codon_start=1

/product="PAM COOH-terminal interactor protein 2"
 /protein_id="AAC53031.2"
 /db_xref="GI:5821768"
 /translation="MAGSGCAMGAPRPFILAFRLKMOVSLRSGSSASVYRVCAG
 TPGSPFALAKOPLPRTTGAAASAEYFRKRRARLEQLQGRINVTLYGVTHFSP
 NVPSCLLELDVSVSELVYSSHQGSSMMIQLCADVDLEALFLHBEGVHADLK
 PNIIIMSAENEFKLIIDPGLSPKGNQVRYIQTGYRAPEALONCIAQLQSDTE
 CTSADVLSIGIILLEMFGMKLKHVTSQEWKANSALIHIFASKVAVNAIPAYH
 LMDLKSMHDEVPSPRIPAEMLCSPPFSIPAPPHIEDLVMLPTVLRLLVDDYL
 ENDEVEDVEDKEKCEKXGPGVSLVLPKRNPGQYFVEYANNGDSAAQKLTLGR
 MFDGFVAATFPISAYRKGYLYTLL"
 ORIGIN
 Query Match 87.8%; Score 1106.4; DB 10; Length 3244;
 Best local Similarity 92.4%; Pred. No. 2,1e-272;
 Matches 1164; Conservative 0; Mismatches 96; Indels 0; Gaps 0;
 QY 1 ATGGCGGAGATCCGGCTGCGCTCGGCGCGAGCCGCGGCTTTCTGAGAGCTTCGGG 60
 Db 162 ATGGCGGAGATCCGGCTGCGCTCGGCGCGAGCCGCGGCTTCCTGAGAGCTTCGGG 221
 QY 61 CGGCTGTGGCAGTACAGAGCCGTCTGGGTACGGCTCTCCGCTCGGTGTATGGGTT 120
 Db 222 CGGCTGTGGCAGTACAGAGCCGTCTGGGTACGGCTCTCCGCTCGGTGTATGGGTT 281
 QY 121 CGCTGTGGCGCAACCTTGAGTCTGCGCCCGCGGCGCTTCAAGAGTCTTTCGCGCAGGA 180
 Db 282 CGCTGTGGCGCAACCTTGAGTCTGCGCCCGCGGCGCTTCAAGAGTCTTTCGCGCAGGA 341
 QY 181 ACCACCGGGGCTGCGGCTCTGCGCGCGAGATGATGATGATGATGATGATGATGATGAT 240
 Db 342 ACCACCGGGGCTGCGGCTCTGCGCGCGAGATGATGATGATGATGATGATGATGATGATGAT 401
 QY 241 GAACAGTTGACAGGTTCACAGAAACATCGTGAATTTGTATGAGTGTATTCATTCACACTT 300
 Db 402 GAGCAGTTGACAGGTTCACAGAAACATCGTGAATTTGTATGAGTGTATTCATTCACACTT 461
 QY 301 TCTCCAAATGTCACATCAAGCTGTCTGTTGCTTGAACCTTGATGTGAGTGTTCGGA 360
 Db 462 TCTCCAAATGTCACATCAAGCTGTCTGTTGCTTGAACCTTGATGTGAGTGTTCGGA 521
 QY 361 TTGCTCTTATATTCAGTACCAAGGATGTTTCATGTGATGATGATGATGATGATGATGATGAT 420
 Db 522 TTGCTCTTATATTCAGTACCAAGGATGTTTCATGTGATGATGATGATGATGATGATGATGAT 581
 QY 421 GATGTTTGGAGCCCTTCTTCTTATCATGAGGCTATGTCATGCGGACCTCAAA 480
 Db 582 GATGTTTGGAGCCCTTCTTCTTATCATGAGGCTATGTCATGCGGACCTCAAA 641
 QY 481 CCACGTAACTATGTTGGAGTGCAGAGATGATGATGATGATGATGATGATGATGATGATGAT 540
 Db 642 CCACGTAACTATGTTGGAGTGCAGAGATGATGATGATGATGATGATGATGATGATGATGAT 701
 QY 541 ACCTTCAAGAGGATCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 Db 702 ACCTTCAAGAGGATCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 761
 QY 601 GAAGCAAAATGCAAAATTTGCTTGGCCCAAGCTGCGCTGCGAGAGTATCAGATGATGATGAT 660
 Db 762 GAAGCAAAATGCAAAATTTGCTTGGCCCAAGCTGCGCTGCGAGAGTATCAGATGATGATGAT 821
 QY 661 TCAGCGTTGATCTGTGAGGCTTAGAATCAATTTACAGAAATGTTCTCGAGAAATGAAA 720
 Db 822 TCAGCGTTGATCTGTGAGGCTTAGAATCAATTTACAGAAATGTTCTCGAGAAATGAAA 881
 QY 721 CTGAACAATACAGTACAGTCTCAGAGATGAGAGCAACAGTTCTGTCTATTATGATGATC 780
 Db 882 CTGAACAATACAGTACAGTCTCAGAGATGAGAGCAACAGTTCTGTCTATTATGATGATC 941
 QY 781 AATTTTGCAGTAAAGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
 Db 942 AATTTTGCAGTAAAGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1001

QY 841 ATCAAAAGCATGCTCATGATGATCCAGAGAAAGTAATCTGTAAGATGGCATTTGTC 900
DB 1002 ATCAAAAGCATGCTCATGATGATCCAGAGAAAGTAATCTGTAAGATGGCATTTGTC 1061
QY 901 AGCCCATTTCTTATGCAATCTCTTTGGCCCTCATATTTGAAGATCTGTCATGCTTCCACT 960
DB 1062 AGCCCATTTCTTATGCAATCTCTTTGGCCCTCATATTTGAAGATCTGTCATGCTTCCACT 1121
QY 961 CCAGTCTAAGATGCTGTAATGCTGATGATGATTTATCTTGGAGATGAAGAATAT 1020
DB 1122 CCAGTCTAAGATGCTGTAATGCTGATGATGATTTATCTTGGAGATGAAGAATAT 1181
QY 1021 GAAGATGTTGTAGAGATGTAAGAGAGAGATGTCAAAATATGAGACCAAGTATCTTA 1080
DB 1182 GAAGATGTTGTAGAGATGTAAGAGAGAGATGTCAAAATATGAGACCAAGTATCTTA 1241
QY 1081 CTGTGTTCCAAAGAAATCTGTCAGAGACCAAGTCTTTGTATGATGAAATGCTGCT 1140
DB 1242 CTGTGTTCCAAAGAAATCTGTCAGAGACCAAGTCTTTGTATGATGAAATGCTGCT 1301
QY 1141 GATTCCAAAGCTGCGAGAAATTAATGATGATGAAAGATGTTGATGGAAGTTGTTGTG 1200
DB 1302 GATTCCAAAGCTGCGAGAAATTAATGATGATGAAAGATGTTGATGGAAGTTGTTGTG 1361
QY 1201 GCTACATTTACCCGCTGAGTGCCTACAGAGGGGATATCTGTATCAAACTTGCTTTAA 1260
DB 1362 GCTACATTTACCCGCTGAGTGCCTACAGAGGGGATATCTGTATCAAACTTGCTTTAA 1421

RESULT 13
MMPKIS 1254 bp mRNA linear ROD 08-JUL-2002
LOCUS Mus musculus partial mRNA for KIS protein kinase (Kis gene).
DEFINITION Y10725.2 GI:21726712
ACCESSION Y10725.2 GI:21726712
VERSION KIS gene; protein kinase.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1254)
Maucuer, A., Ozon, S., Manceau, V., Gavet, O., Lawler, S., Curmi, P. and
Sobel, A.
KIS is a protein kinase with an RNA recognition motif
J. Biol. Chem. 272 (37), 23151-23156 (1997)
92435279
9287318

REFERENCE
AUTHORS Maucuer, A.
TITLE Direct Submission
JOURNAL Submitted (23-JAN-1997) Maucuer A., INSERM U440, U440, 17 Rue du
Fer Moulin, Paris, 75005, FRANCE
REMARK revised by: [4]
3 (bases 1 to 1254)
Maucuer, A.
TITLE Direct Submission
JOURNAL Submitted (02-JUL-2002) A. Maucuer, INSERM U440, U440, 17 Rue du
Fer Moulin, Paris, 75005, FRANCE
COMMENT On Jul 10, 2002 this sequence version replaced gi:1806129.
Overlaps with related sequences X82320, X10725.
FEATURES
source
1. 1254
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="K2"
/tissue_type="brain"
/clone_lib="lambda gt10"
/dev_stage="neonate"
1. 1254
/gene="Kis"
11. >1254
/gene="Kis"

/codon_start=1
/product="KIS protein kinase"
/protein_id="CAA71714.2"
/db_xref="GI:21726713"
/db_xref="GOA:P97343"
/db_xref="Swiss-Prot:P97343"
/translation="NAGSGAMGAPPEPFLAEPLRMLQVQSHLSSGSSAYVRACCG
TPGSPGALKQPLPGTTGAAASAEYFRERKALAEQLQGRNLTLYGVFTTHFS
NPSRCLLELLELDVSVSELLYSHQSGSMWIQHCARDVLEALAEFLHHEGVADLK
PNTILMSAENECFKLIDPGLSPENODVRYIOTDGYAPAEALQNLQAGLOSDTE
CTSAVDIMSLIGTILLEMPSGMKLKHTVYSOEKMSASAIIDHIFASRVVNAIIPAYH
LRDLKSLMLHEDVPEKRIIPAEMLCSPPFSIIPRAPIEDLVNIPPVATLAVLDDYL
ENEDYEIDVEDVEKEECQYGPVSLVLPKENPGRGVFVEYANAGDSKAOKLITGR
MEDGRFVAATFPELSAYRGLY"

ORIGIN

Query Match 87.2%; Score 1098.4; DB 10; Length 1254;
Best Local Similarity 92.7%; Pred. No. 1.9e-270;
Matches 1153; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
QY 1 ATGCGCGGATCCGGCTGCGCTGCGGCGCGAGCCGCGCTTTCTGAGAGCCTTCGCG 60
DB 11 ATGCGCGGATCCGGCTGCGCTGCGGCGCGAGCCGCGCTTTCTGAGAGCCTTCGCG 70
QY 61 CGGCTGTGCGAGGTACAGAGCCGTCTGAGTACGGCTCTCCGCTCGGTGATGGGTT 120
DB 71 CGGCTGTGCGAGGTACAGAGCCGTCTGAGTACGGCTCTCCGCTCGGTGATGGGTT 130
QY 121 CGGCTGTGCGAGGTACAGAGCCGTCTGAGTACGGCTCTCCGCTCGGTGATGGGTT 180
DB 131 CGGCTGTGCGAGGTACAGAGCCGTCTGAGTACGGCTCTCCGCTCGGTGATGGGTT 190
QY 181 ACCACCGGGGCTGCGGCTCTGCTCCGCGAGTATGTTTCGGAAGAGAGGCGGCGCTG 240
DB 191 ACCACCGGGGCTGCGGCTCTGCTCCGCGAGTATGTTTCGGAAGAGAGGCGGCGCTG 250
QY 241 GAACAGTTGCGAGGTACAGAAACATGCTGATCTTGTATGAGTGTTCATCAACTTT 300
DB 251 GAGCAGTTGCGAGGTACAGAAACATGCTGATCTTGTATGAGTGTTCATCAACTTT 310
QY 301 TCTCCAAATGTCATCAAGCTGTCTGTTGTTGAATCTCTGAGATGTCAGTGTTCGGA 360
DB 311 TCTCCAAATGTCATCAAGCTGTCTGTTGTTGAATCTCTGAGATGTCAGTGTTCGGA 370
QY 361 TTGCTCTTATTTCCAGTACCAAGGTTGTTCCATGTGATGATACAGCATTCGCCCA 420
DB 371 TTGCTCTTATTTCCAGTACCAAGGTTGTTCCATGTGATGATACAGCATTCGCCCA 430
QY 421 GATGTTTGGAGCCCTTCTTCTTCATCATGAGGCTATGTCATGCGGACTCAAA 480
DB 431 GATGTTTGGAGCCCTTCTTCTTCATCATGAGGCTATGTCATGCGGACTCAAA 490
QY 481 CCACGTAATATTTGAGAGTGCAGAGATGAATGTTTAACTATGACTTGGACTT 540
DB 491 CCACGTAATATTTGAGAGTGCAGAGATGAATGTTTAACTATGACTTGGACTT 550
QY 541 AGCTTCAAGAGGCAATCAGATGTAAGTATTTATTCAGACAGAGGATTCGCGCTCA 600
DB 551 AGCTTCAAGAGGCAATCAGATGTAAGTATTTATTCAGACAGAGGATTCGCGCTCA 610
QY 601 GAAGCAGAAATTCGAAATGCTTGGCCGAGGCTGCGCTGAGATGATCAGATGATAC 660
DB 611 GAAGCAGAAATTCGAAATGCTTGGCCGAGGCTGCGCTGAGATGATCAGATGATAC 670
QY 661 TCAAGCTGTGATCTGTGAGAGCTTGAAGATATTTATCTGGAATGTTCTCAGGAATGA 720
DB 671 TCAAGCTGTGATCTGTGAGAGCTTGAAGATATTTATCTGGAATGTTCTCAGGAATGA 730
QY 721 CTGAAGATACAGTACAGATCTCAGAGATGAAGAGCAACAGTTCGTCTTATTTATGATCA 780
DB 731 CTGAAGATACAGTACAGATCTCAGAGATGAAGAGCAACAGTTCGTCTTATTTATGATCA 790
QY 781 ATATTTGCCAGTAAAGCAGTGTGATGATGCCGCAATTCAGCCTATCACTTAAGAGACTT 840

Db	791	ATATTTGGCACTAAAGCAATGATGTAATGCCCAATTCAGCCCTACACTCAGAGACTT	850
Qy	841	ATCAAAAGCATGCTTCATGATGATCCAGCAGAAAGATTCCTGCTGAATGGCATTTGTGC	900
Db	851	ATCAAAAGCATGCTTCATGATGACCCGGGGCAGAAAGATCCCTGCTGATGATGGCATTTGTGC	910
Qy	901	AGCCCATCTCTTAAAGATCTCTTTTGGCCCTCATATTTGAAGATCTGTGCTTCCACT	960
Db	911	AGCCCATCTCTTAAAGATCTCTTTTGGCCCTCATATTTGAAGATCTGTGATCTTCCGACT	970
Qy	961	CCAGTGCCTAAGCTGCTGAATGTGTGTGGATCATATGATTTCTTGGGAATGAAGAGGAATAT	1020
Db	971	CCAGTGCCTAAGCTTCTCAATGTGTGTGATATGATTACTTTGAATAAGAAAGATGAAATAT	1030
Qy	1021	GAAGATGTTGTAGAAGATGTATAAAGAGAGTGTCAAAATATATGCAACAGTGTATCTCTTA	1080
Db	1031	GAAGATGTTGTAGAAGATGTATAAAGAGAGTGTCAAAATATATGCAACAGTGTATCTCTTG	1090
Qy	1081	CTGTGTTCCAAAGAGAAATCTGTGCAGAGAGACAAGTCTTTGTTAGATATGCAAAATGTCTGT	1140
Db	1091	CTGTGTTCCAAAGAGAAATCTGTGCAGAGAGACAAGTCTTGTGATATGCGCAACGCTGT	1150
Qy	1141	GATTCCAAAGCTGGCGCAGAAATTACTGACTGGAAGAGATGTTTGATGGGAAGTTGTGTG	1200
Db	1151	GATTCCAAAGCTGCTCAAGATGTCTGACTGGGAGAGATGTTTGAACGGGAAGTTGTGTGTG	1210
Qy	1201	GCTACATTTCTACCCGCTGAGTGCCATCAAGAGGGGATATCTGTA	1244
Db	1211	GCTACATTTCTACCCGCTGAGTGCCATCAAGAGGGGATATCTTTTA	1254

RESULT 14	LOCUS	BC026046	1367 bp	mRNA	linear	PRI 30-JUN-2004
	LOCUS	BC026046	1367 bp	mRNA	linear	PRI 30-JUN-2004
	DESCRIPTION	Homo sapiens kinase interacting with leukemia-associated gene				

ACCESSION	BC026046	GI:19684094
VERSION	BC026046.1	
KEYWORDS	MG.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	

REFERENCE
1 (bases 1 to 1367)
AUTHORS
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16699-16903 (2002)

JOURNAL
PUBLISHED
REFERENCE
AUTHORS
TITLE
JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477992

2 (bases 1 to 1367)

Strausberg, R.

Direct Submission

Submitted (20-MAR-2002)

National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer

REMARK : NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT : Contact: MGC help desk
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcgc.bc.ca
 Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
 Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
 Featherstone, Malachi Griffith, Obi Griffith, Ran Gini, Nancy Liao,
 Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
 Tekla Olson, Diana Palmquist, Anca Petrescu, Anna Lisa Prabhu,
 Parvaneh Saeedi, J.R. Santos, Angélique Scherch, Ursula Skalska,
 Duane Smalls, Jeff Stoltz, Miranda Tsai, George Yang, Jacque
 Schein, Asim Siddiqui, Rob Holt, Marco Marra.

```

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/INL at: http://image.inl.gov
Series: IRAL plate: 42 Row: f Column: 6
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi.: 21385438.
location/Qualifiers
1. .1367

```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon.9606"
/clone="MGC:39391 IMAGE:4649261"
/issue_type="Eye, retinoblastoma"
/clone_id="NH_MGC_16"
/lab_host="DH10B-R"
/note="vector: pOT87"
1..1367
/gene="KIS"
/locus_tag="KIS"
/locus_id="LOC100506481"
/locus_name="KIS"
/locus_path="12p11.23"
/locus_size="1367"
/locus_type="mRNA"
/locus_xref="MIM:103480"
/locus_yref="12p11.23"
/locus_zref="12p11.23"
/locus_zref2="12p11.23"
/locus_zref3="12p11.23"
/locus_zref4="12p11.23"
/locus_zref5="12p11.23"
/locus_zref6="12p11.23"
/locus_zref7="12p11.23"
/locus_zref8="12p11.23"
/locus_zref9="12p11.23"
/locus_zref10="12p11.23"
/locus_zref11="12p11.23"
/locus_zref12="12p11.23"
/locus_zref13="12p11.23"
/locus_zref14="12p11.23"
/locus_zref15="12p11.23"
/locus_zref16="12p11.23"
/locus_zref17="12p11.23"
/locus_zref18="12p11.23"
/locus_zref19="12p11.23"
/locus_zref20="12p11.23"
/locus_zref21="12p11.23"
/locus_zref22="12p11.23"
/locus_zref23="12p11.23"
/locus_zref24="12p11.23"
/locus_zref25="12p11.23"
/locus_zref26="12p11.23"
/locus_zref27="12p11.23"
/locus_zref28="12p11.23"
/locus_zref29="12p11.23"
/locus_zref30="12p11.23"
/locus_zref31="12p11.23"
/locus_zref32="12p11.23"
/locus_zref33="12p11.23"
/locus_zref34="12p11.23"
/locus_zref35="12p11.23"
/locus_zref36="12p11.23"
/locus_zref37="12p11.23"
/locus_zref38="12p11.23"
/locus_zref39="12p11.23"
/locus_zref40="12p11.23"
/locus_zref41="12p11.23"
/locus_zref42="12p11.23"
/locus_zref43="12p11.23"
/locus_zref44="12p11.23"
/locus_zref45="12p11.23"
/locus_zref46="12p11.23"
/locus_zref47="12p11.23"
/locus_zref48="12p11.23"
/locus_zref49="12p11.23"
/locus_zref50="12p11.23"
/locus_zref51="12p11.23"
/locus_zref52="12p11.23"
/locus_zref53="12p11.23"
/locus_zref54="12p11.23"
/locus_zref55="12p11.23"
/locus_zref56="12p11.23"
/locus_zref57="12p11.23"
/locus_zref58="12p11.23"
/locus_zref59="12p11.23"
/locus_zref60="12p11.23"
/locus_zref61="12p11.23"
/locus_zref62="12p11.23"
/locus_zref63="12p11.23"
/locus_zref64="12p11.23"
/locus_zref65="12p11.23"
/locus_zref66="12p11.23"
/locus_zref67="12p11.23"
/locus_zref68="12p11.23"
/locus_zref69="12p11.23"
/locus_zref70="12p11.23"
/locus_zref71="12p11.23"
/locus_zref72="12p11.23"
/locus_zref73="12p11.23"
/locus_zref74="12p11.23"
/locus_zref75="12p11.23"
/locus_zref76="12p11.23"
/locus_zref77="12p11.23"
/locus_zref78="12p11.23"
/locus_zref79="12p11.23"
/locus_zref80="12p11.23"
/locus_zref81="12p11.23"
/locus_zref82="12p11.23"
/locus_zref83="12p11.23"
/locus_zref84="12p11.23"
/locus_zref85="12p11.23"
/locus_zref86="12p11.23"
/locus_zref87="12p11.23"
/locus_zref88="12p11.23"
/locus_zref89="12p11.23"
/locus_zref90="12p11.23"
/locus_zref91="12p11.23"
/locus_zref92="12p11.23"
/locus_zref93="12p11.23"
/locus_zref94="12p11.23"
/locus_zref95="12p11.23"
/locus_zref96="12p11.23"
/locus_zref97="12p11.23"
/locus_zref98="12p11.23"
/locus_zref99="12p11.23"
/locus_zref100="12p11.23"
/locus_zref101="12p11.23"
/locus_zref102="12p11.23"
/locus_zref103="12p11.23"
/locus_zref104="12p11.23"
/locus_zref105="12p11.23"
/locus_zref106="12p11.23"
/locus_zref107="12p11.23"
/locus_zref108="12p11.23"
/locus_zref109="12p11.23"
/locus_zref110="12p11.23"
/locus_zref111="12p11.23"
/locus_zref112="12p11.23"
/locus_zref113="12p11.23"
/locus_zref114="12p11.23"
/locus_zref115="12p11.23"
/locus_zref116="12p11.23"
/locus_zref117="12p11.23"
/locus_zref118="12p11.23"
/locus_zref119="12p11.23"
/locus_zref120="12p11.23"
/locus_zref121="12p11.23"
/locus_zref122="12p11.23"
/locus_zref123="12p11.23"
/locus_zref124="12p11.23"
/locus_zref125="12p11.23"
/locus_zref126="12p11.23"
/locus_zref127="12p11.23"
/locus_zref128="12p11.23"
/locus_zref129="12p11.23"
/locus_zref130="12p11.23"
/locus_zref131="12p11.23"
/locus_zref132="12p11.23"
/locus_zref133="12p11.23"
/locus_zref134="12p11.23"
/locus_zref135="12p11.23"
/locus_zref136="12p11.23"
/locus_zref137="12p11.23"
/locus_zref138="12p11.23"
/locus_zref139="12p11.23"
/locus_zref140="12p11.23"
/locus_zref141="12p11.23"
/locus_zref142="12p11.23"
/locus_zref143="12p11.23"
/locus_zref144="12p11.23"
/locus_zref145="12p11.23"
/locus_zref146="12p11.23"
/locus_zref147="12p11.23"
/locus_zref148="12p11.23"
/locus_zref149="12p11.23"
/locus_zref150="12p11.23"
/locus_zref151="12p11.23"
/locus_zref152="12p11.23"
/locus_zref153="12p11.23"
/locus_zref154="12p11.23"
/locus_zref155="12p11.23"
/locus_zref156="12p11.23"
/locus_zref157="12p11.23"
/locus_zref158="12p11.23"
/locus_zref159="12p11.23"
/locus_zref160="12p11.23"
/locus_zref161="12p11.23"
/locus_zref162="12p11.23"
/locus_zref163="12p11.23"
/locus_zref164="12p11.23"
/locus_zref165="12p11.23"
/locus_zref166="12p11.23"
/locus_zref167="12p11.23"
/locus_zref168="12p11.23"
/locus_zref169="12p11.23"
/locus_zref170="12p11.23"
/locus_zref171="12p11.23"
/locus_zref172="12p11.23"
/locus_zref173="12p11.23"
/locus_zref174="12p11.23"
/locus_zref175="12p11.23"
/locus_zref176="12p11.23"
/locus_zref177="12p11.23"
/locus_zref178="12p11.23"
/locus_zref179="12p11.23"
/locus_zref180="12p11.23"
/locus_zref181="12p11.23"
/locus_zref182="12p11.23"
/locus_zref183="12p11.23"
/locus_zref184="12p11.23"
/locus_zref185="12p11.23"
/locus_zref186="12p11.23"
/locus_zref187="12p11.23"
/locus_zref188="12p11.23"
/locus_zref189="12p11.23"
/locus_zref190="12p11.23"
/locus_zref191="12p11.23"
/locus_zref192="12p11.23"
/locus_zref193="12p11.23"
/locus_zref194="12p11.23"
/locus_zref195="12p11.23"
/locus_zref196="12p11.23"
/locus_zref197="12p11.23"
/locus_zref198="12p11.23"
/locus_zref199="12p11.23"
/locus_zref200="12p11.23"
/locus_zref201="12p11.23"
/locus_zref202="12p11.23"
/locus_zref203="12p11.23"
/locus_zref204="12p11.23"
/locus_zref205="12p11.23"
/locus_zref206="12p11.23"
/locus_zref207="12p11.23"
/locus_zref208="12p11.23"
/locus_zref209="12p11.23"
/locus_zref210="12p11.2
```

```

/gene="KIS"
/codon_start=1
/product="kinase interacting stathmin"
/protein_id="AAH26046.1"
/db_xref="GI:19684095"
/db_xref="LocustID:127933"
/translacion="MAGSGCAGWGAEPFRFLFAGRLTQWQVSRLLGSSGSASVYRVRCGG
NPGSPGALIKOFLPPGTTLGAASAQSGYFKERKALDLOLQGHNIIVTLVGYFTLIFHE
NPSPCLLELDVSVSELLIYSSHOCSMMWIDHCARDVLEALAFILHIEGVYADILK
PRNLTGASBNBCEGLIDFGSLPKRGNQDVXYIOTDGYRAPEALQNTCAOGLQSDPHE
CTSAVDVMSYCIITLLEMFSGMKLKHTRVRSQWKNSSATDILFASAVVAALPAYHY
LRDLIKSMHDDPSRRIPALMALCSPPFSIFPAHIEDLVMLPTPVLRILNVLDDVDYIL
ENESEYEGLC"

```

	Query Match	Similarity	84.6%;	Score 1065.6;	DB 9;	Length 1367;
	Best Local	Similarity	92.6%;	Pred. No. 5.1e-262;		
	Matches 1167;	Conservative	0;	Mismatches 4;	Indels 89;	Gaps 1;
QY	1	ATGAGCGGANTCGGCTGCGCTGGGAGCGCGGAGCCGTCGCTTTTCTGAGAGCCCTTTGCGG	60			
DB	159	ATGAGCGGAGATCGGCTGCGCTGGGAGCGGAGCGCGCTTTCTGAGAGCCCTTTGCGG	218			
QY	61	CGGCTGTGGCAGGTACAAGAGCCGTTGGGGTATGAGGCTCTCCGCTCGGATTCGAGTT	120			
DB	219	CGGCTGTGGCAGGTACAAGAGCCGTTGGGGTATGAGGCTCTCCGCTCGGATTCGAGTT	278			
QY	121	CGTCTCTCGGCAACCTTGAGCTCGCCGCCCGGCGCCCTCAGGCAAGTTCTTCCGCGCAGA	180			
DB	279	CGTCTCTCGGCAACCTTGAGCTCGCCGCCCGGCGCCCTCAGGCAAGTTCTTCCGCGCAGA	338			
QY	181	ACCAACCGGGGCTGGCGGCTCTGCGCGCGAGTATGTTTCCGCAAGAAGAGGGCGGGCGCTG	240			
DB	339	ACCAACCGGGGCTGGCGGCTCTGCGCGCGAGTATGTTTCCGCAAGAAGAGGGCGGGCGCTG	398			


```

QY 241 GAA CAG TTG CAG GGT C A C A G A A A C A T G T G A C T T T G T A T G A G T G T T T A C A A T C C A C T T T 300
Db 399 G A A C A G T T G C A G G G T C A C A G A A A C A T G T G A C T T T G T A T G A G T G T T T A C A A T C C A C T T T 458
QY 301 T C T C C A A T T G C C A T A C G C T G T C T T T G C T T G A A C T C C T G A A T G T C A G T T T T G G A A 360
Db 459 T C T C C A A T T G C C A T A C G C T G T C T T T G A A C T C C T G A A T G T C A G T T T T G G A A 518
QY 361 T T G C T C T A T T T C C A G T C A C A G G G T G T T C A T G A T G A A T G A A T G A A T G C G G C C G A 420
Db 519 T T G C T C T A T T T C C A G T C A C A G G G T G T T C A T G A T G A A T G A A T G A A T G C G G C C G A 578
QY 421 G A T G T T T G G A G G C C C T T G C T T T C T T C A T C A T A G A G G G C T A T G C A T G C G A C C T C A A 480
Db 579 G A T G T T T G A G G C C C T T G C T T T C T T C A T C A T A G A G G G C T A T G C A T G C G A C C T C A A 638
QY 481 C C A C G T A C A T A T T T G T G A G T G C A G A G A A T G T T T A A C T C A T T G A C T T T G A C T T 540
Db 639 C C A C G T A C A T A T T T G T G A G T G C A G A G A A T G T T T A A C T C A T T G A C T T T G A C T T 698
QY 541 A G C T T C A A A G A G G C A T C A G A T G T A A G T A T T C A G A C A G G G T A T G G G G C T C C A 600
Db 699 A G C T T C A A A G A G G C A T C A G A T G T A A G T A T T C A G A C A G G G T A T G G G G C T C C A 758
QY 601 G A A C A A A T T G C A A A A T T G C T T G C C C A G C T G C C T G C A G A G T A T C A G A A T G T A C C 660
Db 759 G A A G C A A A T T G C A A A A T T G C T T G C C C A G C T G C C T G C A G A T G A T T C A G A A T G T A C C 818
QY 661 T C A G C T T G A T C T G T G A G C C T T A G A A T C A T T T A C T G A A A T G T T C T C A G A A T G A A 720
Db 819 T C A G C T T G A T C T G T G A G C C T T A G A A T C A T T T T A C T G A A A T G T T C T C A G A A T G A A 878
QY 721 C T A A A C A T A C A G T C A G A T C T C A G A A T G A A G C A A A G T C T G C T A T T A T T A T T A T C A C 780
Db 879 C T A A A C A T A C A G T C A G A T C T C A G A A T G A A G C A A A G T C T G C T A T T A T T A T T A T C A C 938
QY 781 A T A T T T G C C A T A A G A G A G T G A A T G C C G A A T T C A G C C T A T T A C C T A A G A G A C C T T 840
Db 939 A T A T T T G C C A T A A G A G A G T G A A T G C C G A A T T C A G C C T A T T A C C T A A G A G A C C T T 998
QY 841 A T A A A A A G C A T G C T T C A T G A T T C C A G A C A A A A A T T C T G C T G A A A T G S C A T T T G C 900
Db 999 A T C A A A A G C A T G C T T C A T G A T G A T C C A G A C A A A A A T T C T G C T G A A A T G S C A T T T G C 1058
QY 901 A G C C C A T T C T T A G C A T T C C T T T G C C C C A T A T T A A G A T G S C A T T G C C C A C T 960
Db 1059 A G C C C A T T C T T A G C A T T C C T T T T G C C C C A T A T T A A G A T G S C A T T G C C C A C T 1118
QY 961 C C A G T C T A A G A C T G C T G A A T G T C T G A T G A T A T T C T T G G A A T G A A G A A T A T 1020
Db 1119 C C A G T C T A A G A C T G C T G A A T G T C T G A T G A T A T T C T T G A A T G A A G A A T A T 1178
QY 1021 G A A G A T G T T A G A A T G T A A A A G A G A G T C A A A A A T A T A T G A C C A G T G A T C T A 1080
Db 1179 G A A G ----- 1182
QY 1081 C T T G T T C A A A G A A A T C C T G G C A G A G A C A G T C T T T G T G A T G A A T G C T G G T 1140
Db 1183 ----- G T C T T G T G A T G A T G C A A A T G C T G G T 1209
QY 1141 G A T T C A A A G C T G C G A G A A A T T A C T G A C T G A A G A T G T T G A T G G A A G T T G T T G 1200
Db 1210 G A T T C A A A G C T G C G A G A A A T T A C T G A C T G A A G A T G T T G A T G G A A G T T G T T G 1269
QY 1201 G C T A C A T T T A C C C G C T G A G T G C T C A A G A G G G A T A T C T G T A T A A A C T T G C T T A A 1260
Db 1270 G C T A C A T T T A C C C G C T G A G T G C T C A A G A G G G A T A T C T G T A T A A A C T T G C T T A A 1329

```

```

RESULT 15
AK058195
LOCUS AK058195 1135 bp mRNA linear PRI 12-SEP-2003

```

```

DEFINITION Homo sapiens cDNA FLJ25466 f18, clone TST09304, highly similar to
ACCESSION AK058195
VERSION AK058195.1 GI:16554273
KEYWORDS oligo capping, f18 (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Suzuki, O., Sasaki, N., Aotsuka, S., Shoji, T., Ichihara, T.,
Shiohata, N., Matsumoto, K., Hirano, M., Sano, S., Nomura, R.,
Yoshikawa, Y., Matsumura, Y., Moriya, S., Chiba, E., Momiyama, H.,
Onogawa, S., Kaeriyama, S., Satoh, N., Matsunawa, H., Takahashi, E.,
Katoka, R., Kuga, N., Kuroda, A., Satoh, I., Kamata, K., Takami, S.,
Terashima, Y., Watanabe, M., Suzuki, Y., Hata, H., Nakagawa, K.,
Mizuno, S., Morinaga, M., Kawamura, M., Sugiyama, T., Irie, R.,
Otsuki, T., Sato, H., Nishikawa, T., Sugiyama, A., Kawakami, B.,
Nagai, K., Isogai, T. and Sugano, S.
NEO human cDNA sequencing project
TITLE Unpublished
JOURNAL 2 (bases 1 to 1135)
AUTHORS Sugano, S. and Suzuki, Y.
TITLE Direct Submission
JOURNAL Submitted (24-OCT-2001) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure, Human Genome
Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail: f18@med.ms.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT NEO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan, cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction and 5'-end one pass sequencing: Institute of Medical
Science, University of Tokyo, Laboratory of Genome Structure, Human
Genome Center; 3'-end one pass sequencing: RAB; clone selection for
full insert sequencing: RAB and Helix Research Institute.
FEATURES
source 1..1135
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TST09304"
/tissue_type="testis"
/clone_lib="TST"
/note="Cloning vector: pME18SFL3"
ORIGIN
Query Match 78.6%; Score 989.8; DB 9; Length 1135;
Best Local Similarity 99.8%; Pred. No. 1.3e-242;
Matches 991; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 268 G G A C C T T T G A T G A G A G T G T T A A C A A T C A C T T T C T C C A A A T G A G C C A C A G C T G T G 327
Db 123 G T A C C T T G T A T G A G A G T T T A C A A T C A C A C T T T C T C C A A A T G A G C C A C A G C T G T G 182
QY 328 T T G C T T G A A C T C T G A T G T C A G T G T T T C G G A A T T G C T T A T A T T C A G T C A C A G G G T 387
Db 183 T T G C T T G A A C T C T G A T G T C A G T G T T T C G G A A T T G C T T A T A T T C A G T C A C A G G G T 242
QY 388 T G T T C A A T G T G A T A C A G A T T G C C C C G A G A T G T T T G A A G G C C T T G C T T T C T T 447
Db 243 T G T T C A A T G T G A T A C A G A T T G C C C G A G A T G T T T G A A G G C C T T G C T T T C T T 302
QY 448 C A T C A T A G A G G C T A T G T C A T G C G G A C C T C A A C A C G T A C A T A T T G T G A G T G C A G 507
Db 303 C A T C A T A G A G G C T A T G T C A T G C G G A C C T C A A C A C G T A C A T A T T G T G A G T G C A G 362
QY 508 A A T G A A T G T T T A A C T A T T G A C T T T G A C T T A G C T T C A A A G A A G G C A A T C A G A A T G A 567
Db 363 A A T G A A T G T T T A A C T A T T G A C T T T G A C T T A G C T T C A A A G A A G G C A A T C A G A A T G A 422
QY 568 A A G T A T T C A G A C A G A C G G G T A T C G G G C T C A A A G C A A G A A T T G C T T G G C C 627

```

Db 423 AACTATATTCAGACAGACGGGTATCGGGCTCCAGAAAGAGAAATTCGAAAATTCCTTGCC 482
Qy 628 CAGGCTGGCCTGAGAGTGAATACAGAAATGACTCAGCTGTTGATCTGTGGAGCCTAGGA 687
Db 483 CAGGCTGGCCTGAGAGTGAATACAGAAATGACTCAGCTGTTGATCTGTGGAGCCTAGGA 542
Qy 688 ATCATTTTACTGGAATGTTCTCAGAAATGAACTGAAACATACAGTCAAGATCTCAGGAA 747
Db 543 ATCATTTTACTGGAATGTTCTCAGAAATGAACTGAAACATACAGTCAAGATCTCAGGAA 602
Qy 748 TGGAAAGCAAAACAGTCTGCTATTTATTTGATCAGATATTTGCGAGTAAAGCATGTGTAAT 807
Db 603 TGGAAAGCAAAACAGTCTGCTATTTATTTGATCAGATATTTGCGAGTAAAGCATGTGTAAT 662
Qy 808 GCGGCAATTCAGACCTATCACTTAAGAGACCTTATCAAAAGCATGCTTCAATGATCCA 867
Db 663 GCGGCAATTCAGACCTATCACTTAAGAGACCTTATCAAAAGCATGCTTCAATGATCCA 722
Qy 868 AGCAGAGAATTCCTGCTGAAATGCGATGTCAGGCCATTCCTTTAGCATTCCTTTGCG 927
Db 723 AGCAGAGAATTCCTGCTGAAATGCGATGTCAGGCCATTCCTTTAGCATTCCTTTGCG 782
Qy 928 CCTCATATTGAAGATCTGTCATGCTTCCCATCTCAGTGTAAAGCTGTGAATGTGCTG 987
Db 783 CCTCATATTGAAGATCTGTCATGCTTCCCATCTCAGTGTAAAGCTGTGAATGTGCTG 842
Qy 988 GATGATGATTAATCTTGGAATGAAGAGAAATGAAAGATGTTGTAAGATGTAAAGAG 1047
Db 843 GATGATGATTAATCTTGGAATGAAGAGAAATGAAAGATGTTGTAAGATGTAAAGAG 902
Qy 1048 GAGTGTCAAAAATATGGAACAGAGTGTATCTCTACTTGTCCAAAGAAAATCTGGCAGA 1107
Db 903 GAGTGTCAAAAATATGGAACAGAGTGTATCTCTACTTGTCCAAAGAAAATCTGGCAGA 962
Qy 1108 GGAACAAGCTTTGTTGATGCAAAATGCTGTGATTCGAAAGCTGCGAGAAATTAATG 1167
Db 963 GGAACAAGCTTTGTTGATGCAAAATGCTGTGATTCGAAAGCTGCGAGAAATTAATG 1022
Qy 1168 ACTGGAAGGATGTTTGAATGGGAAGTTTGTGNGGCTACATTCACCCGCTGAGTGCCTAC 1227
Db 1023 ACTGGAAGGATGTTTGAATGGGAAGTTTGTGNGGCTACATTCACCCGCTGAGTGCCTAC 1082
Qy 1228 AAGAGGGGATATCTGTATCAAACTTGCTTTAA 1260
Db 1083 AAGAGGGGATATCTGTATCAAACTTGCTTTAA 1115

Search completed: November 29, 2004, 14:49:20
Job time : 5514 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2004, 11:23:25 ; Search time 653.5 Seconds
(without alignments)
10121.301 Million cell updates/sec

Title: US-10-798-532-3

Perfect score: 1260
Sequence: 1 atggggggatccgctgcgc.....tgatcaacctgtcttaa 1260

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 413486 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1260	100.0	1260	3	AAZ51356 Human txa
2	1258.4	99.9	1260	3	AAZ51355 Human wil
3	1256.8	99.7	2008	6	ABQ93183 Human cod
4	1255.2	99.6	1260	4	AAI66828 Human pro
5	1255.2	99.6	2622	4	AAI66828 Human pro
6	1255.2	99.6	4065	4	AAH72714 Human cer
7	1255.2	99.6	4065	5	ADL45238 Human ova
8	1253.6	99.5	1260	4	AA506734 Polynucle
9	1253.6	99.5	1824	4	AAAD18818 Human kin
10	1173.4	93.1	3998	12	ADQ24401 Human sof
11	1106.4	87.8	3244	10	ADBS2857 Primary r
12	458.6	36.4	529	4	AAH70764 Human cer
13	445.8	35.4	722	3	AAA02535 Human cer
14	426.6	33.9	470	4	AAH71966 Human cer
15	375.4	29.8	461	4	AAH69101 Human cer
16	272.2	21.6	411	6	ABQ92225 Human ORF
17	203.8	16.2	231	3	AAA01259 Human ORF
18	177.4	14.1	419	5	AD173740 Human ova
19	177.4	14.1	419	5	AD167345 Human ova
20	176.2	14.0	546	6	ABQ51127 Oligonuc1
21	176.2	14.0	546	6	ABQ51126 Oligonuc1

C	22	176	14.0	547	6	ABQ44465	Abq44465 Oligonuc1
	23	176	14.0	547	6	ABQ44464	Abq44464 Oligonuc1
	24	157.2	12.5	546	6	ABQ53129	Abq53129 Oligonuc1
C	25	157.2	12.5	546	6	ABQ53128	Abq53128 Oligonuc1
	26	157.2	12.5	547	6	ABQ44463	Abq44463 Oligonuc1
C	27	157.2	12.5	547	6	ABQ44462	Abq44462 Oligonuc1
	28	148.4	11.8	442	9	ACH28557	ACH28557 Human adu
	29	85	6.7	198	5	ADJ38977	ADJ38977 Human ova
	30	63.4	5.0	65	6	ABN52650	ABN52650 Mouse bpl
	31	63.2	5.0	549	10	ADK58517	ADK58517 Plant DNA
	32	57.8	4.6	1196	12	ADJ39614	ADJ39614 Plant CDN
	33	53.8	4.3	1364	12	ADJ42189	ADJ42189 Plant CDN
	34	47.4	3.8	1747	6	AA594774	AA594774 Human DNA
	35	47.4	3.8	1863	3	AAZ47544	AAZ47544 Human pro
	36	47.4	3.8	2044	8	ABX70932	ABX70932 Novel hum
	37	47.4	3.8	2887	12	ADJ95566	ADJ95566 Human Dyr
	38	46.8	3.7	7565	12	ADK70703	ADK70703 Murine CR
	39	45.4	3.6	1581	6	ABR88905	ABR88905 Murine md
	40	44.4	3.5	2658	3	AAA70213	AAA70213 Plasmodiu
	41	43.8	3.5	1584	6	ABL58336	ABL58336 Rat YAK1
	42	43	3.4	681	10	ADK52877	ADK52877 Plant DNA
C	43	42.8	3.4	600	12	ADJ44008	ADJ44008 Plant CDN
	44	42.4	3.4	1497	3	AAA70154	AAA70154 Plasmodiu
	45	42.2	3.3	2095	3	AA36941	AA36941 Arabidops

ALIGNMENTS

RESULT 1	AAZ51356	AAZ51356 standard; DNA; 1260 BP.
ID	AAZ51356	
XX	AAZ51356	
AC	AAZ51356	
XX	06-JUN-2000	(first entry)
DT		
XX		
DE		Human transdominant mutant serine/threonine kinase KIS (hKIS) gene.
XX		
KW	KIS; human; hKIS; serine/threonine kinase; cell proliferation; G1 phase;	
KW	transdominant gene; mutant; cyclin-dependent kinase inhibitor; CKI; p27;	
KW	modulator; treatment; cell proliferative disease; vascular disorder;	
KW	gene therapy; atherosclerosis; restenosis; ds.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	1..1260
FT		/*tag= a
FT		/product= "Human transdominant mutant KIS protein"
FT		/transl_except= (pos:556..561, Asn)
FT		161
FT	misc_feature	/*tag= b
FT		/note= "A to G substitution resulting in a Lys to Arg
FT		change in the amino acid sequence"
XX		
XX	WO200011165-A1.	
XX		
PD	02-MAR-2000.	
XX		
PF	20-AUG-1999.	99WO-US018903.
XX		
PR	21-AUG-1998.	98US-0097710P.
XX		
PA	(NABE/) NABEL G J.	
PA	(NABE/) NABEL B G.	
XX		
PI	Nabel GJ, Nabel EG;	
XX		
XX	WPI; 2000-237648/20.	
DR	P-FSDB; AAY70306.	
XX		
PT	Novel serine/threonine kinase hKIS polynucleotides and polypeptides used	

for inhibiting the cyclin kinase inhibitor p27, and so alter cell proliferation.

Claim 7; Page 59; 70pp; English.

The present DNA sequence encodes a transdominant mutant human KIS (hKIS), constructed by site directed mutagenesis. A single nucleotide substitution (A to G) results in a lysine to arginine change in the protein sequence. hKIS is a serine/threonine kinase, that acts as an inhibitory kinase of cyclin-dependent kinase inhibitor (CKI), p27. hKIS controls cell proliferation and is localised predominantly in the nucleus. It binds to CKI p27 and inhibits its ability to arrest cells in G1 phase. The hKIS sequences are used to modulate cell proliferation and treat cell proliferative and vascular diseases. The polynucleotide sequence may be used in gene therapy to treat vascular disorders such as restenosis or atherosclerosis

Sequence 1260 BP; 310 A; 280 C; 336 G; 334 T; 0 U; 0 Other;

Query Match 100.0%; Score 1260; DB 3; Length 1260;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGGCGGAGATCCGGCTGGCCCTGGGGCGCGAAGCCCGCGTTTCTGAGAGCGCTTCGGG 60
DB 1 ATGGCGGAGATCCGGCTGGCCCTGGGGCGCGAAGCCCGCGTTTCTGAGAGCGCTTCGGG 60
QY 61 CGGCTGTGAGGATGACAGAGCCGCTGTGGTAGCGGCTCCCGCTCGGATGATCCGGATT 120
DB 61 CGGCTGTGAGGATGACAGAGCCGCTGTGGTAGCGGCTCCCGCTCGGATGATCCGGATT 120
QY 121 CGGCTGTGCGGCAACCTTGCTGCGCCCGCGCGCCCTCAAGAGATTCTTGGCGCAGAA 180
DB 121 CGGCTGTGCGGCAACCTTGCTGCGCCCGCGCGCCCTCAAGAGATTCTTGGCGCAGAA 180
QY 181 ACCACCGGGGCTGGCGGCTCTGGCGCGAGATGTTTCCGAAAGAGAGGGCGGCGT 240
DB 181 ACCACCGGGGCTGGCGGCTCTGGCGCGAGATGTTTCCGAAAGAGAGGGCGGCGT 240
QY 241 GAACAGTTGACAGGGTCAACAGAAATCATGTAATCTTGTATGAGTGTTCATCACTTT 300
DB 241 GAACAGTTGACAGGGTCAACAGAAATCATGTAATCTTGTATGAGTGTTCATCACTTT 300
QY 301 TCTCCAAATGTGCGATCAGCTGTCTGTGCTTGAATCTCTGATGATCAAGTGTTCGAA 360
DB 301 TCTCCAAATGTGCGATCAGCTGTCTGTGCTTGAATCTCTGATGATCAAGTGTTCGAA 360
QY 361 TTGCTCTTATATCCAGTCAACGAGGTTGTTCAATGATGATTAAGCAATTCGCGCGA 420
DB 361 TTGCTCTTATATCCAGTCAACGAGGTTGTTCAATGATGATTAAGCAATTCGCGCGA 420
QY 421 GATGTTTGGAGGCCCTTGCTTTCATCATGAGGGCTATGCTCATGCGACCTCAAA 480
DB 421 GATGTTTGGAGGCCCTTGCTTTCATCATGAGGGCTATGCTCATGCGACCTCAAA 480
QY 481 CCACGTAACATATTTGAGAGTGCAGAGATGAATGTTTAACTCATGACTTGAAGCTT 540
DB 481 CCACGTAACATATTTGAGAGTGCAGAGATGAATGTTTAACTCATGACTTGAAGCTT 540
QY 541 AGCTTAAAGAGAGCATCAGATGATTAATTCAGACAGCGGATAGCGGCTCCA 600
DB 541 AGCTTAAAGAGAGCATCAGATGATTAATTCAGACAGCGGATAGCGGCTCCA 600
QY 601 GAAGCAGAAATGCAAAATTCCTTGCCAGAGCTGGCTGCAGAGATACAGAAATGACC 660
DB 601 GAAGCAGAAATGCAAAATTCCTTGCCAGAGCTGGCTGCAGAGATACAGAAATGACC 660
QY 661 TCAGCTGTGATGTGAGAGCTTGAAGATCATTTTAACTGAAATGTTTTCAGAAATGAA 720
DB 661 TCAGCTGTGATGTGAGAGCTTGAAGATCATTTTAACTGAAATGTTTTCAGAAATGAA 720
QY 721 CTGAACATACATCATGATCTCAGAAATGAGCAAAAGTCTGCTATATTAATGATC 780
DB 721 CTGAACATACATCATGATCTCAGAAATGAGCAAAAGTCTGCTATATTAATGATC 780

```

```

DB 721 CTGAACATACATCATGATCTCAGAAATGAGCAAAAGTCTGCTATTAATGATC 780
QY 781 ATATTGGCCAGTAAGCAGATGATGATGCGCAATTCAGCTTACCTAAGAGACCTT 840
DB 781 ATATTGGCCAGTAAGCAGATGATGATGCGCAATTCAGCTTACCTAAGAGACCTT 840
QY 841 ATCAAAGAGATCTTCATGATGATATCCAGCAGAAATTCCTGCTGAAATGCAATTCG 900
DB 841 ATCAAAGAGATCTTCATGATGATATCCAGCAGAAATTCCTGCTGAAATGCAATTCG 900
QY 901 ACCCATCTTTAGGATTCCTTTGCCCCATATTTGAAGATCTGTCATGCTTCCACT 960
DB 901 ACCCATCTTTAGGATTCCTTTGCCCCATATTTGAAGATCTGTCATGCTTCCACT 960
QY 961 CCAAGTCTAAGATCTGTAATGCTGTAAGATGATATCTTGGAAATGAAGAGATAT 1020
DB 961 CCAAGTCTAAGATCTGTAATGCTGTAAGATGATATCTTGGAAATGAAGAGATAT 1020
QY 1021 GAAGATGTTTGAAGATGTAAGAGAGATGTAAGAGATGTAAGAGATGTAAGATGTA 1080
DB 1021 GAAGATGTTTGAAGATGTAAGAGAGATGTAAGAGATGTAAGAGATGTAAGATGTA 1080
QY 1081 CTTGTTCCAAAGAAATCTCTGCGAGAGCAAGTCTTGTGATGATGCAAAATGCTG 1140
DB 1081 CTTGTTCCAAAGAAATCTCTGCGAGAGCAAGTCTTGTGATGATGCAAAATGCTG 1140
QY 1141 GATTCCAAAGCTGCGCAGAAATTAAGTGAAGATGTTTGAAGATGTTTGTG 1200
DB 1141 GATTCCAAAGCTGCGCAGAAATTAAGTGAAGATGTTTGAAGATGTTTGTG 1200
QY 1201 GCTACATTTACCGCTGAGTCCCTCAAGAGGAGATCTGTATCAAACTTGTCTTA 1260
DB 1201 GCTACATTTACCGCTGAGTCCCTCAAGAGGAGATCTGTATCAAACTTGTCTTA 1260

RESULT 2
AAZ51355
ID AAZ51355 standard; DNA; 1260 BP.
AC AAZ51355;
DT 06-JUN-2000 (first entry)
DE Human wild type serine/threonine kinase KIS (hKIS) gene.
XX KIS; human; hKIS; serine/threonine kinase; cell proliferation; G1 phase;
XX dominant gene; cyclin-dependent kinase inhibitor; CKI; p27; modulator;
XX treatment; cell proliferative disease; vascular disorder; gene therapy;
XX atherosclerosis; restenosis; ds.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH 1. 1260
FT CDS /*tag= a
FT /product= "Human wild type KIS protein"
FT /transl_except= (pos:556..561, Aaa)
XX
XX WO20001165-A1.
XX
XX 02-MAR-2000.
XX
XX 20-AUG-1999; 99WO-US018903.
XX
XX 21-AUG-1998; 98US-0097710P.
XX
XX (NABE/) NABEL G J.
XX (NABE/) NABEL E G.
XX
XX Nabel GJ, Nabel EG;
XX
XX WPI: 2000-237648/20.
XX P-PSDB; AAT70305.
DR

```

XX Novel serine/threonine kinase hKIS polynucleotides and polypeptides used
PT for inhibiting the cyclin kinase inhibitor p27, and so alter cell
PT proliferation.

PS Claim 1; Page 56; 70pp; English.

CC The present DNA sequence encodes a wild type human KIS (hKIS), isolated
CC from a human B-cell library, using a yeast two hybrid screening system.
CC hKIS is a serine/threonine kinase, that acts as an inhibitory kinase of
CC cyclin-dependent kinase inhibitor (CKI), p27. hKIS is a dominant gene
CC that controls cell proliferation and is localised predominantly in the
CC nucleus. It binds to CKI p27 and inhibits its ability to arrest cells in
CC G1 phase. It has 99% homology to rat serine/threonine protein kinase KIS.
CC The hKIS sequences are used to modulate cell proliferation and treat cell
CC proliferative and vascular diseases. The polynucleotide sequence may be
CC used in gene therapy to treat vascular disorders such as restenosis or
CC atherosclerosis

CC Sequence 1260 BP; 311 A; 280 C; 335 G; 334 T; 0 U; 0 Other;

CC Query Match 99.9%; Score 1258.4; DB 3; Length 1260;

CC Best Local Similarity 99.9%; Pred. No. 0;

CC Matches 1259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGGGGATCCGGGCTGGCGGCGGAGCGCGCGGCTTTCTGAGAGGCTTCGGG 60
DB 1 ATGGCGGGGATCCGGGCTGGCGGCGGAGCGCGCGGCTTTCTGAGAGGCTTCGGG 60
QY 61 CGGCTGTGGCAGGTACAGAGCCGCTGTGGGTAGCGGCTCCGCTCGGTGTATCGGGTT 120
DB 61 CGGCTGTGGCAGGTACAGAGCCGCTGTGGGTAGCGGCTCCGCTCGGTGTATCGGGTT 120
QY 121 CGGTGTGGCGCAACCTGCTGCCCCCGCGCGCTTCAAGCATGTTCTGCGCAGGA 180
DB 121 CGGTGTGGCGCAACCTGCTGCCCCCGCGCGCTTCAAGCATGTTCTGCGCAGGA 180
QY 181 ACCACCGGGGCTGGCGGCTCTGCGCGAGTATGTTTCCGAAAGAGGGGCGGCTG 240
DB 181 ACCACCGGGGCTGGCGGCTCTGCGCGAGTATGTTTCCGAAAGAGGGGCGGCTG 240
QY 241 GAACAGTTGACGGGTACAGAAACATCGTGACTTTGTATGAGTGTTCACATCACTTT 300
DB 241 GAACAGTTGACGGGTACAGAAACATCGTGACTTTGTATGAGTGTTCACATCACTTT 300
QY 301 TCTCCAAATGTGCAATCAAGCTGTCTGTGCTTGAACCTCTGATGTCAGTGTTCGAA 360
DB 301 TCTCCAAATGTGCAATCAAGCTGTCTGTGCTTGAACCTCTGATGTCAGTGTTCGAA 360
QY 361 TTGCTCTTATATTCAGTCAACGAGGTGTTCATGTGATGATTAACGATTCGCGCCA 420
DB 361 TTGCTCTTATATTCAGTCAACGAGGTGTTCATGTGATGATTAACGATTCGCGCCA 420
QY 421 GATGTTTGGAGGCGCTTCTTCTTCATCATGAGGGCTATATCCATGCGGACTCAAA 480
DB 421 GATGTTTGGAGGCGCTTCTTCTTCATCATGAGGGCTATATCCATGCGGACTCAAA 480
QY 481 CCAAGTAACTATTTGAGAGTGAAGATGATTTTAACTCAATTCATTGGAATT 540
DB 481 CCAAGTAACTATTTGAGAGTGAAGATGATTTTAACTCAATTCATTGGAATT 540
QY 541 AGCTTCAAAAGAGCATGAGATGTAATATATTCAGCAACGGGTATCGGGCTCCA 600
DB 541 AGCTTCAAAAGAGCATGAGATGTAATATATTCAGCAACGGGTATCGGGCTCCA 600
QY 601 GAAGCAATTTGCAAAATTCCTTGCCAGAGCTGCGCTGAGAGTATACAGATGATCC 660
DB 601 GAAGCAATTTGCAAAATTCCTTGCCAGAGCTGCGCTGAGAGTATACAGATGATCC 660
QY 661 TCAGCTGTGATCTGTGAGCCTTAGAATCATTTTACGAAATGTTCTCAGAAATGAAA 720
DB 661 TCAGCTGTGATCTGTGAGCCTTAGAATCATTTTACGAAATGTTCTCAGAAATGAAA 720

QY 721 CTGAACATACAGTCAGATCTCAGGAATGAAAGCAACAGTCTCTATTTATGATCAG 780
DB 721 CTGAACATACAGTCAGATCTCAGGAATGAAAGCAACAGTCTCTATTTATGATCAG 780
QY 781 ATATTGGCAGTAAAGCATGCTGTAATCCCGCAATTCAGGCTATCACTTAAGACCTT 840
DB 781 ATATTGGCAGTAAAGCATGCTGTAATCCCGCAATTCAGGCTATCACTTAAGACCTT 840
QY 841 ATCAAAAGCATGCTCATGATGATCCAGAGAAATTCCTGCTGAATGAGCATTTGTC 900
DB 841 ATCAAAAGCATGCTCATGATGATCCAGAGAAATTCCTGCTGAATGAGCATTTGTC 900
QY 901 AGCCCATTTCTTGAATTCCTTTTCCCTCATATTAAGATCTGTGATGCTTCCACT 960
DB 901 AGCCCATTTCTTGAATTCCTTTTCCCTCATATTAAGATCTGTGATGCTTCCACT 960
QY 961 CCAAGTCTAAGCTCTGTAATGCTGTGATGATGATTTATCTTTGGAAATGAAGAAAT 1020
DB 961 CCAAGTCTAAGCTCTGTAATGCTGTGATGATGATTTATCTTTGGAAATGAAGAAAT 1020
QY 1021 GAAGATGTTGTAAGATGTAAGAGAGTCTCAAAATATAGACAGTGTATCTCTA 1080
DB 1021 GAAGATGTTGTAAGATGTAAGAGAGTCTCAAAATATAGACAGTGTATCTCTA 1080
QY 1081 CTGTGTTCCAAAGAAATCTCTGCGCAGAGCAAGTCTTGTGATGATGCAATGCTGCT 1140
DB 1081 CTGTGTTCCAAAGAAATCTCTGCGCAGAGCAAGTCTTGTGATGATGCAATGCTGCT 1140
QY 1141 GATTCCAAAGCTGCGCAAGAAATTAATCTGACTGGAAGATGTTTGAATGGAAGTTTGTG 1200
DB 1141 GATTCCAAAGCTGCGCAAGAAATTAATCTGACTGGAAGATGTTTGAATGGAAGTTTGTG 1200
QY 1201 GCTAATCTTACCGCTGAGTGCCTTCAAGAGGGGATATCTGTATCAAACTTGCTTAA 1260
DB 1201 GCTAATCTTACCGCTGAGTGCCTTCAAGAGGGGATATCTGTATCAAACTTGCTTAA 1260

RESULT 3

ABQ9383
ID ABQ9383 standard; cDNA; 2008 BP.

XX ABQ9383;

DT 25-FEB-2003 (first entry)

XX Human coding sequence SEQ ID 116.

XX Human; expressed sequence tag; EST; haematopoietic disorder;

KM central nervous system disease; viral infection;

KM peripheral nervous system disease; non-healing wound; infectious disease;

KM immune deficiency; immune disorder; bacterial infection; allergic; cancer;

KM fungal infection; autoimmune disorder; coagulation disorder; noctropic;

KM antiallergic; antiinflammatory; immunosuppressive; neuroprotective;

KM cytostatic; hemostatic; virucide; antibacterial; fungicide;

KM immunostimulant; cerebroprotective; gene therapy; gene; ss.

XX Homo sapiens.

OS Homo sapiens.

PN WO200259260-A2.

XX 01-AUG-2002.

XX 16-NOV-2001; 2001WO-US042950.

XX 17-NOV-2000; 2000US-00714936.

XX (HYSE-) HYSEQ INC.

PI Tang YF, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
PI Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX MPI; 2002-590824/63.
DR N-PSDB; ABP64797.

XX New isolated polynucleotide, useful in research, diagnostic or
 PT therapeutic methods, e.g. preventing or treating disorders involving
 PT aberrant protein expression or biological activity.

PS Claim 1, SEQ ID NO 116; 394pp; English.

XX The present invention relates to novel human coding sequences (AB093268-
 CC AB093608) and proteins (ABP64682-ABP65022). The sequences are useful in
 CC therapeutic, diagnostic and research methods. The polynucleotides may be
 CC used in the field of molecular biology as hybridisation probes, primers
 CC for PCR, for chromosome and gene mapping, for the recombinant production
 CC of protein, or in generation of anti-sense DNA or RNA. The
 CC polynucleotides are useful in diagnostics as expressed sequence tags
 CC (ESTs) for identifying expressed genes or for physical mapping of the
 CC human genome. The proteins may be used as molecular weight markers, or as
 CC nutritional sources or supplements. The proteins may be used to maintain
 CC and expand cell population in a totipotent or pluripotent state
 CC useful for re-engineering damaged or diseased tissues, transplantation,
 CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The
 CC polynucleotides and proteins are useful for preventing, treating or
 CC ameliorating disorders involving aberrant protein expression or
 CC biological activity, e.g. haematopoietic disorders, central/peripheral
 CC nervous system diseases, mechanical and traumatic disorders, non-healing
 CC wounds, immune deficiencies and disorders, infectious diseases caused by
 CC viral, bacterial or fungal infection, autoimmune disorders, allergic
 CC reactions and conditions, coagulation disorders, or cancer. The
 CC polynucleotide sequences of the invention were assembled from ESTs
 CC isolated mainly by sequencing by hybridisation, and in some cases,
 CC sequences obtained from one or more public databases. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPD at
 CC ftp.wipo.int/pub/published_pct_sequences

SO Sequence 2008 BP; 491 A; 451 C; 488 G; 578 T; 0 U; 0 Other;

Query Match 99.7%; Score 1256.8; DB 6; Length 2008;

Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1258; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCGGAGATCCGGCTCGGCGCGGAGAGCGCGCTTTTGAGAGCGCTTCGGG 60
 DB 79 ATGGCGGAGATCCGGCTCGGCGCGGAGAGCGCGCTTTTGAGAGCGCTTCGGG 138
 QY 61 CGGCTGTGCGAGGTACAGAGCGCTCTGGTAGCGGCTCCGCTCGGTATCGGGTT 120
 DB 139 CGGCTGTGCGAGGTACAGAGCGCTCTGGTAGCGGCTCCGCTCGGTATCGGGTT 198
 QY 121 CGGCTGTGCGAGGTACAGAGCGCTCCGCGCGGCGCTTCAGAGGATTTTCCGCGAGGA 180
 DB 199 CGGCTGTGCGAGGTACAGAGCGCTCCGCGCGGCGCTTCAGAGGATTTTCCGCGAGGA 258
 QY 181 ACCACCGGGGCTCGGCGCTCTGCGCGGAGTATGTTTCCGAAAGAGAGGGCGGCTG 240
 DB 259 ACCACCGGGGCTCGGCGCTCTGCGCGGAGTATGTTTCCGAAAGAGAGGGCGGCTG 318
 QY 241 GAACAGTTGCGAGGTACAGAGAAATCATGACTTTGATGAGAGTGTAAATCCACTTT 300
 DB 319 GAACAGTTGCGAGGTACAGAGAAATCATGACTTTGATGAGAGTGTAAATCCACTTT 378
 QY 301 TCTCCAAATGTCAGTACAGCTGTCTGTGTTGAAATCTCGTGAATGTCAGTGTTCGAA 360
 DB 379 TCTCCAAATGTCAGTACAGCTGTCTGTGTTGAAATCTCGTGAATGTCAGTGTTCGAA 438
 QY 361 TTGCTCTTATATCCAGTACAGAGGGTGTTCATGATGATGATAGAGATTGCGCGGA 420
 DB 439 TTGCTCTTATATCCAGTACAGAGGGTGTTCATGATGATGATAGAGATTGCGCGGA 498
 QY 421 GATGTTTGGAGGCGCTTGTCTTTTTCATCATGAGGGCTATGTCAGTCCGAGCTCAAA 480
 DB 499 GATGTTTGGAGGCGCTTGTCTTTTTCATCATGAGGGCTATGTCAGTCCGAGCTCAAA 558
 QY 481 CCACTTAACATATTTGAGAGTGCAGAGAAATGTTTAACTCATGACTTTGAGACTT 540

DB 559 CCACCTAACATATTTGAGAGTGCAGAGATGAAATGTTTAACTCATGACTTTGAGACTT 618
 QY 541 AGCTTCAAGAGGAGCATCAGATGTAAAGTATTTAGACAGAGCGGTATCGGCTCCA 600
 DB 619 AGCTTCAAGAGGAGCATCAGATGTAAAGTATTTAGACAGAGCGGTATCGGCTCCA 678
 QY 601 GAAGCAGATTTGCAAAATTTGCTTGCGCCAGGCTGCGCTGAGAGTATACAGATGATAC 660
 DB 679 GAAGCAGATTTGCAAAATTTGCTTGCGCCAGGCTGCGCTGAGAGTATACAGATGATAC 738
 QY 661 TAGCTGTGATCTGTGAGAGCTTAGAGATCATTTTACTGGAATTTTCTCAGAAATGAAA 720
 DB 739 TAGCTGTGATCTGTGAGAGCTTAGAGATCATTTTACTGGAATTTTCTCAGAAATGAAA 798
 QY 721 CTGAAACATACAGTACAGATCTCAGAGATGGAAGGCAACAGTTCTGATTTATTTAGAC 780
 DB 799 CTGAAACATACAGTACAGATCTCAGAGATGGAAGGCAACAGTTCTGATTTATTTAGAC 858
 QY 781 ATATTGTCAGTAAAGCAGTGTGTAATCCGCAATTCAGGCTATCAGCTTAAGAGACCTT 840
 DB 859 ATATTGTCAGTAAAGCAGTGTGTAATCCGCAATTCAGGCTATCAGCTTAAGAGACCTT 918
 QY 841 ATCAAAGCATGCTTCATGATGATTCACAGCAGAAATTTCTGCTGAATGCAATTTGTC 900
 DB 919 ATCAAAGCATGCTTCATGATGATTCACAGCAGAAATTTCTGCTGAATGCAATTTGTC 978
 QY 901 AGCCATTTCTTAGCATTCCTTTTSCCCTCATATTGAAGATCTGTGATGCTTCCACT 960
 DB 979 AGCCATTTCTTAGCATTCCTTTTSCCCTCATATTGAAGATCTGTGATGCTTCCACT 1038
 QY 961 CCAAGTCTAAGACTGTGTAATGCTGTGATGATGATTTCTTGGGAATGAAGAGAAATAT 1020
 DB 1039 CCAAGTCTAAGACTGTGTAATGCTGTGATGATGATTTCTTGGGAATGAAGAGAAATAT 1098
 QY 1021 GAAGATGTTTGAAGATGTAAAGAGAGTCTCAAAATATGACACAGTGTATCTCTA 1080
 DB 1099 GAAGATGTTTGAAGATGTAAAGAGAGTCTCAAAATATGACACAGTGTATCTCTA 1158
 QY 1081 CTTGTTCCAAAGAAATCTTGCGCAGAGACCAAGTCTTTGTTAGATGCAAAATCTGCT 1140
 DB 1159 CTTGTTCCAAAGAAATCTTGCGCAGAGACCAAGTCTTTGTTAGATGCAAAATCTGCT 1218
 QY 1141 GATTCAAAGCTGCGCAGAAATTAATCTGAGAGATGTTTATGGAAGTGTGTG 1200
 DB 1219 GATTCAAAGCTGCGCAGAAATTAATCTGAGAGATGTTTATGGAAGTGTGTG 1278
 QY 1201 GCTACATTTACCCGCTGAGTCCCTACAGAGGGATATCTGTATCAAACTTGCTTTAA 1260
 DB 1279 GCTACATTTACCCGCTGAGTCCCTACAGAGGGATATCTGTATCAAACTTGCTTTAA 1338

RESULT 4
 AA166829
 ID AA166829 standard; cDNA; 1260 BP.
 XX
 AC AA166829;
 XX
 DT 07-JAN-2002 (first entry)
 XX
 XX Human protein kinase polypeptide 13887 coding sequence.
 DE
 XX Protein kinase; 3714; 16742; 23546; 13887; cancer; bone disorder; human;
 KW cytosolic; antiinflammatory; immunosuppressive; cardiac; hepatotropic;
 KM gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1260
 FT /tag= a
 FT /product= "protein kinase 13887"
 XX

PN MO200173050-A2.
 XX 04-OCT-2001.
 XX 23-MAR-2001; 2001MO-US009483.
 XX 24-MAR-2000; 2000US-0191846P.
 XX (MILL-) MILLENNIUM PHARM INC..
 PA
 XX
 PI Meyers R;
 XX
 DR WPI; 2001-611632/70.
 XX P-PSDB; AAC65767.
 PT
 PT New human protein kinase polypeptides, 3714, 16742, 23546 and 13887,
 PT useful in diagnosis of cancer or cellular proliferation or
 PT differentiation disorders and to screen for polypeptide modulators useful
 PT to treat such conditions.
 XX
 PS Claim 1; Fig 10A-C; 16pp; English.
 CC The invention provides novel human protein kinase polypeptides, 3714,
 CC 16742, 23546 and 13887 and nucleic acid molecules encoding them. The
 CC protein kinase polypeptides can be expressed by standard recombinant
 CC methodology. 3714, 16742, 23546 or 13887 nucleic acids and polypeptides
 CC are useful for diagnostic and screening methods to identify subjects (at
 CC risk of) having cancer or cellular proliferation and/or differentiation
 CC disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and
 CC modulators are useful for the treatment of cancer, particularly colon
 CC cancer or cellular proliferation and/or differentiation disorders. Other
 CC disorders associated with 3714, 16742, 23546 or 13887 expression or
 CC activity that can be treated include bone related disorders, inflammatory
 CC disorders, autoimmune diseases, cardiovascular disorders and liver
 CC diseases. The present sequence represents a human protein kinase
 CC polypeptide 13887 coding sequence
 CC
 SQ Sequence 1260 BP; 312 A; 279 C; 334 G; 335 T; 0 U; 0 Other;
 Query Match 99.6%; Score 1255.2; DB 4; Length 1260;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1257; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 ATGGCGGAGATCCGCGCTGCGCGTGGGCGGAGCGCGCGCTTCTGAGGCGCTTCGCG 60
 DB 1 ATGGCGGAGATCCGCGCTGCGCGTGGGCGGAGCGCGCGCTTCTGAGGCGCTTCGCG 60
 QY 61 CGGCTGTGGCAGGTACAGAGCCGCTGTGGTAGCGGCTCTCCGCTCGGTGTATCGG 120
 DB 61 CGGCTGTGGCAGGTACAGAGCCGCTGTGGTAGCGGCTCTCCGCTCGGTGTATCGG 120
 QY 121 CGGTGTGGGCGCAACCTGTGCTGCCCGCGCGCTTCAAGCATGCTTGGCGGAGA 180
 DB 121 CGGTGTGGGCGCAACCTGTGCTGCCCGCGCGCTTCAAGCATGCTTGGCGGAGA 180
 QY 181 ACCACCGGGGCTGCGGCTCTGCCGCGAGTATGTTTCCGAAAGAGAGGCGGCGCTG 240
 DB 181 ACCACCGGGGCTGCGGCTCTGCCGCGAGTATGTTTCCGAAAGAGAGGCGGCGCTG 240
 QY 241 GAACAGTTGAGGGTCAAGAAACATGTAATCTTGTATGAGTGTGTTCAATCACTTT 300
 DB 241 GAACAGTTGAGGGTCAAGAAACATGTAATCTTGTATGAGTGTGTTCAATCACTTT 300
 QY 301 TCTCCAAAGTGGCATCAGCTGTCTGTGTTGAACTCTGTGATGTGAGTGTTCGAA 360
 DB 301 TCTCCAAAGTGGCATCAGCTGTCTGTGTTGAACTCTGTGATGTGAGTGTTCGAA 360
 QY 361 TTGCTCTTATATTCAGAGTCAAGGAGTGTTCATGTGATGATGATGATGATGATG 420
 DB 361 TTGCTCTTATATTCAGAGTCAAGGAGTGTTCATGTGATGATGATGATGATGATG 420
 QY 421 GATGTTTGGAGGCCCTTGCTTCTTCTCATCATGAGGCGTATGTCCATGCGGACTTAA 480
 DB 421 GATGTTTGGAGGCCCTTGCTTCTTCTCATCATGAGGCGTATGTCCATGCGGACTTAA 480

DB 421 GATGTTTGGAGGCCCTTGCTTCTTCTCATCATGAGGCGTATGTCCATGCGGACTTAA 480
 QY 481 CCACGTAAATATTTGGAGTGCAGAGATGATGATGATGATGATGATGATGATGATGAT 540
 DB 481 CCACGTAAATATTTGGAGTGCAGAGATGATGATGATGATGATGATGATGATGATGAT 540
 QY 541 AGCTTCAAGAGGCAATCAGATGTAAAGTATATTCAGACAGAGGATGATGATGATGAT 600
 DB 541 AGCTTCAAGAGGCAATCAGATGTAAAGTATATTCAGACAGAGGATGATGATGATGAT 600
 QY 601 GAAGCAGAAATTCGAAATTTGCTGGCCGAGGCTGCGCTGAGATGATCAGAAATGATC 660
 DB 601 GAAGCAGAAATTCGAAATTTGCTGGCCGAGGCTGCGCTGAGATGATCAGAAATGATC 660
 QY 661 TCAGCTGTGATCTGTGGAGCCCTAGGAATCATTTTATCGGAATGTTTCAGAAATGAA 720
 DB 661 TCAGCTGTGATCTGTGGAGCCCTAGGAATCATTTTATCGGAATGTTTCAGAAATGAA 720
 QY 721 CTGAACATACAGTCAGATCTCAGAAATGAAAGCAAGTTCTGATTAATGATC 780
 DB 721 CTGAACATACAGTCAGATCTCAGAAATGAAAGCAAGTTCTGATTAATGATC 780
 QY 781 ATATTTCCAGTAAAGCAGTGTGATGTCGCAATTCAGCTTATCAGTAAAGCCTT 840
 DB 781 ATATTTCCAGTAAAGCAGTGTGATGTCGCAATTCAGCTTATCAGTAAAGCCTT 840
 QY 841 ATCAAAAGATCTTCAATATATATCCAGCAAGAAATTCCTGTAATGATGATGATGAT 900
 DB 841 ATCAAAAGATCTTCAATATATATCCAGCAAGAAATTCCTGTAATGATGATGATGAT 900
 QY 901 AGCCCATTTCTTGAATCTCTTTGCGCTCATATGGAATCTGTGATGATGATGATGAT 960
 DB 901 AGCCCATTTCTTGAATCTCTTTGCGCTCATATGGAATCTGTGATGATGATGATGAT 960
 QY 961 CCAAGTCTAAGATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
 DB 961 CCAAGTCTAAGATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
 QY 1021 GAAGATGTTGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
 DB 1021 GAAGATGTTGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
 QY 1081 CTTGTTCCAAAGAAATCTTGGCAGAGCAAGCTTTGTTGATGATGATGATGATGAT 1140
 DB 1081 CTTGTTCCAAAGAAATCTTGGCAGAGCAAGCTTTGTTGATGATGATGATGATGAT 1140
 QY 1141 GATTCGAAAGCTGCGCAGAAATTAAGTGAAGAGATGATGATGATGATGATGATGAT 1200
 DB 1141 GATTCGAAAGCTGCGCAGAAATTAAGTGAAGAGATGATGATGATGATGATGATGAT 1200
 QY 1201 GCTACATTTACCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
 DB 1201 GCTACATTTACCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
 RESULT 5
 AA166828
 ID AA166828 standard; cDNA; 2622 BP.
 XX
 AC AA166828;
 XX
 DT 07-JAN-2002 (first entry)
 XX
 XX Human protein kinase polypeptide 13887 encoding cDNA.
 DE
 XX Protein kinase; 3714; 16742; 23546; 13887; cancer; bone disorder; human;
 KW cytoelastic; antiinflammatory; immunosuppressive; cardiac; hepatotropic;
 KW gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 5'UTR 1..266

FT CDS /*tag= a
FT 267..1529 /*tag= b
FT /product= "protein kinase 13887"
FT /note= "coding sequence specifically claimed"
FT 1530..2622
FT /*tag= c
XX MO200173050-A2.
XX 04-OCT-2001.
XX 23-MAR-2001; 2001MO-US009483.
XX 24-MAR-2000; 2000US-0191846P.
XX (MILL-) MILLENNIUM PHARM INC.
XX Meyers R;
XX MPI; 2001-611632/70.
XX P-PDB; AAG65767.
XX New human protein kinase polypeptides, 3714, 16742, 23546 and 13887,
XX useful in diagnosis of cancer or cellular proliferation or
XX differentiation disorders and to screen for polypeptide modulators useful
XX to treat such conditions.
XX Claim 1; Fig 10A-C; 169pp; English.
XX
XX The invention provides novel human protein kinase polypeptides, 3714,
XX 16742, 23546 and 13887 and nucleic acid molecules encoding them. The
XX protein kinase polypeptides can be expressed by standard recombinant
XX methodology. 3714, 16742, 23546 or 13887 nucleic acids and polypeptides
XX are useful for diagnostic and screening methods to identify subjects (at
XX risk of) having cancer or cellular proliferation and/or differentiation
XX disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and
XX modulators are useful for the treatment of cancer, particularly colon
XX cancer or cellular proliferation and/or differentiation disorders. Other
XX disorders associated with 3714, 16742, 23546 or 13887 expression or
XX activity that can be treated include bone related disorders, inflammatory
XX disorders, autoimmune diseases, cardiovascular disorders and liver
XX diseases. The present sequence represents a human protein kinase
XX polypeptide 13887 encoding cDNA
SQ Sequence 2622 BP; 652 A; 567 C; 627 G; 776 T; 0 U; 0 Other;
Query Match 99.6%; Score 1255.2; DB 4; Length 2622;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1257; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGGCGGATCCCGCTGCGCTGCGGCGCGAGACCGCGCTTTCTGAGAGCTTCGAG 60
DB 267 ATGGCGGATCCCGCTGCGCTGCGGCGCGAGACCGCGCTTTCTGAGAGCTTCGAG 326
QY 61 CGGCTGTGCGAGTACAGAGCCGCTTGGGTAGCGGCTCTCGCTCGGTGTATCGGGTT 120
DB 327 CGGCTGTGCGAGTACAGAGCCGCTTGGGTAGCGGCTCTCGCTCGGTGTATCGGGTT 386
QY 121 CGGCTGTGCGAGACCTGCGCTGCGGCGCGGCGCTTGAAGAGTTCTTGGCGCAGGA 180
DB 387 CGGCTGTGCGAGACCTGCGCTGCGGCGCGGCGCTTGAAGAGTTCTTGGCGCAGGA 446
QY 181 ACCACCGGGGCTCGGCTTGGCGCGAGTATGTTTCCGAAAGAGAGGCGGCGCTG 240
DB 447 ACCACCGGGGCTCGGCTTGGCGCGAGTATGTTTCCGAAAGAGAGGCGGCGCTG 506
QY 241 GAACAGTTGAGGGGTACAGAAACATCTGTGACTTTGTATGAGTGTATTCAATCACTTT 300
DB 507 GAACAGTTGAGGGGTACAGAAACATCTGTGACTTTGTATGAGTGTATTCAATCACTTT 566
QY 301 TCTCCAAATGTGCATCAGCTGTCTGTGTTGAAGTCTGTGATGTCAAGTGTGTTGGAA 360

DB 567 TCTCCAAATGTGCATCAGCTGTCTGTGTTGAAGTCTGTGATGTCAAGTGTGTTGGAA 626
QY 361 TTGCTCTTATATTCAGATCAGCAGGGTGTTCATATGTGAGTATCAGCATTTGCCCGCA 420
DB 627 TTGCTCTTATATTCAGATCAGCAGGGTGTTCATATGTGAGTATCAGCATTTGTGCCGA 686
QY 421 GATGTTTGGAGCCCTTGTCTTTCTTCATCATGAGGCTATGTCCATGCGGACTCAAA 480
DB 687 GATGTTTGGAGCCCTTGTCTTTCTTCATCATGAGGCTATGTCCATGCGGACTCAAA 746
QY 481 CCACGTAAATATTTGGAGTGCAGAGTGAATGTTTAACTGATGACTTTGAGACT 540
DB 747 CCACGTAAATATTTGGAGTGCAGAGTGAATGTTTAACTGATGACTTTGAGACT 806
QY 541 AGCTTCAAGAGGCAATCAGATGTAAGTATATTCAGACAGACGGGTATCGGCTCCA 600
DB 807 AGCTTCAAGAGGCAATCAGATGTAAGTATATTCAGACAGACGGGTATCGGCTCCA 866
QY 601 GAACAGAAATTGCAAAATTTGCTGGCCAGGCTGGCTGCAAGTGTATCAGATGTAC 660
DB 867 GAACAGAAATTGCAAAATTTGCTGGCCAGGCTGGCTGCAAGTGTATCAGATGTAC 926
QY 661 TCAGCTGTGATCTGTGAGGCTAGAGAAATCATTTTACTGGAATGTTCTCAGAAATGAA 720
DB 927 TCAGCTGTGATCTGTGAGGCTAGAGAAATCATTTTACTGGAATGTTCTCAGAAATGAA 986
QY 721 CTGAAACATACAGTCAGATCTCAGAAATGGAAGGCAACAGTTCTGTATTTATGATAC 780
DB 987 CTGAAACATACAGTCAGATCTCAGAAATGGAAGGCAACAGTTCTGTATTTATGATAC 1046
QY 781 ATATTTGCCATGAAAGCAATGCTGTAATGCGCAATTCAGCTTATCATTAAGAACCTT 840
DB 1047 ATATTTGCCATGAAAGCAATGCTGTAATGCGCAATTCAGCTTATCATTAAGAACCTT 1106
QY 841 ATCAAAAGCATCTTCATGATGATCAAGCAGAAAGATTCCTGTGAAATGCAATTTGTC 900
DB 1107 ATCAAAAGCATCTTCATGATGATCAAGCAGAAAGATTCCTGTGAAATGCAATTTGTC 1166
QY 901 AGCCCATCTTTAGCATTCCTTTTCCCTCATATTTGAAGATCTGTCATGCTTCCACT 960
DB 1167 AGCCCATCTTTAGCATTCCTTTTCCCTCATATTTGAAGATCTGTCATGCTTCCACT 1226
QY 961 CCAGTCTTAAGCTGCTGAATGTCTGATGATGATTTCTTGGAAATGGAAGAAAT 1020
DB 1227 CCAGTCTTAAGCTGCTGAATGTCTGATGATGATTTCTTGGAAATGGAAGAAAT 1286
QY 1021 GAAGATGTGTAGAGATGTAAGAGAGTCAAAAATATGACCAAGTGTATCTCTA 1080
DB 1287 GAAGATGTGTAGAGATGTAAGAGAGTCAAAAATATGACCAAGTGTATCTCTA 1346
QY 1081 CTTGTTCCAAAGAAATCTTGGCAGAGCAAGTCTTTTGTGAGTATGCAATGCTGCT 1140
DB 1347 CTTGTTCCAAAGAAATCTTGGCAGAGCAAGTCTTTTGTGAGTATGCAATGCTGCT 1406
QY 1141 GATTCGAAGCTGCGCAAAATTAATGACTGGAAGAGTGTGATGAGGAAGTTTGTG 1200
DB 1407 GATTCGAAGCTGCGCAAAATTAATGACTGGAAGAGTGTGATGAGGAAGTTTGTG 1466
QY 1201 GCTACATTTACCCGCTGAGTCTTACAAAGAGGAGTATCTGTATCAAACTTTGCTTAA 1260
DB 1467 GCTACATTTACCCGCTGAGTCTTACAAAGAGGAGTATCTGTATCAAACTTTGCTTAA 1526
RESULT 6
AAH72714
ID AAH72714 standard; cDNA; 4065 BP.
XX AAH72714;
XX AC AAH72714;
XX XX
DT 19-SEP-2001 (first entry)
XX XX
DE Human cervical cancer marker nucleic acid 3988.

KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX Homo sapiens.
XX MO200142467-A2.
XX
XX
XX 14-JUN-2001.
XX
XX 08-DEC-2000; 2000MO-US033312.
XX
XX 08-DEC-1999; 99US-0169681P.
PR 21-DEC-1999; 99US-0171350P.
PR 14-MAR-2000; 2000US-0189315P.
PR 12-MAY-2000; 2000US-0203791P.
PR 09-JUN-2000; 2000US-0210600P.
PR 21-JUL-2000; 2000US-0220114P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Deeds J, Berger A, Zhao X;
XX WPI; 2001-375006/39.
XX
XX New isolated nucleic acid for diagnosing and treating cervical cancer and
PT for assessing and detecting compounds for treating the cancer.
XX
XX
XX Claim 1; Page 810-811; 1051pp; English.
XX
XX The invention relates to novel genes (AAH68727-AAH7383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy
XX
XX
SQ Sequence 4065 BP; 1068 A; 822 C; 905 G; 1265 T; 0 U; 5 Other;
Query Match 99.6%; Score 1255.2; DB 4; Length 4065;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1257; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGGCGGATTCGCGCTGCGCGTGGGCGGAGCGCGGCTTTCGAGGCGTTCGCG 60
DB 268 ATGGCGGATTCGCGCTGCGCGTGGGCGGAGCGCGCGGCTTTCGAGGCGTTCGCG 327
QY 61 CGGCTGTGGCAGGTACAGAGCCGTCTGAGTACGCGCTCTCCGCTTCGAGTATCGG 120
DB 328 CGGCTGTGGCAGGTACAGAGCCGTCTGAGTACGCGCTCTCCGCTTCGAGTATCGG 387
QY 121 CGGCTGTGGCAGGTACAGAGCCGTCTGAGTACGCGCTCTCCGCTTCGAGTATCGG 180
DB 388 CGGCTGTGGCAGGTACAGAGCCGTCTGAGTACGCGCTCTCCGCTTCGAGTATCGG 447
QY 181 ACCACCGGGGCTGCGCGCTCTGCGCGGATGATGTTCCGCAAGAGGGGCGGCTG 240
DB 448 ACCACCGGGGCTGCGCGCTCTGCGCGGATGATGTTCCGCAAGAGGGGCGGCTG 507
QY 241 GAACAGTTGCAAGGTACAGAGAAACATGTCATCTTGTATGAGAGTTCATCACTTT 300
DB 508 GAACAGTTGCAAGGTACAGAGAAACATGTCATCTTGTATGAGAGTTCATCACTTT 567
QY 301 TCTCCAAAGTGTGTCATGAGGTGTGTTGTTGAACTCTGATGTCAGTGTTCGAA 360
DB 568 TCTCCAAAGTGTGTCATGAGGTGTGTTGTTGAACTCTGATGTCAGTGTTCGAA 627
QY 361 TTGCTCTTATATTCAGTACGAGGTGTGTTGATGATGATGATGATGATGATGATG 420
DB 628 TTGCTCTTATATTCAGTACGAGGTGTGTTGATGATGATGATGATGATGATGATG 687
QY 421 GATGTTTGGAGGCGCTTGTCTTCTTCATCATGAGGCTATGTCATGCGGACTCAAA 480
|||||

DB 688 GATGTTTGGAGGCGCTTGTCTTCTTCATCATGAGGCTATGTCATGCGGACTCAAA 747
QY 481 CCACGTAAATATGTTGAGTGCAGAGATGATGTTTAACTGATGACTTGGACTT 540
DB 748 CCACGTAAATATGTTGAGTGCAGAGATGATGTTTAACTGATGACTTGGACTT 807
QY 541 ACCTTCAAGAGGCAATCAGATGTAAGTATATTCAGACAGAGCGGTATCGGCTCA 600
DB 808 ACCTTCAAGAGGCAATCAGATGTAAGTATATTCAGACAGAGCGGTATCGGCTCA 867
QY 601 GAACAGAAATTCGAAATTCCTGGCCGAGCTGCGCTGTCAGAGTATCAGATGATC 660
DB 868 GAACAGAAATTCGAAATTCCTGGCCGAGCTGCGCTGTCAGAGTATCAGATGATC 927
QY 661 TCAGCTGTGATCTGTGAGCGCTAGGAATCATTTTACGGAATGTTCCAGGAATGAA 720
DB 928 TCAGCTGTGATCTGTGAGCGCTAGGAATCATTTTACGGAATGTTCCAGGAATGAA 987
QY 721 CTGAAACATACAGTCAATCTCAGAGATGGAAGGCAACAGTTCTGATTTATGATCAC 780
DB 988 CTGAAACATACAGTCAATCTCAGAGATGGAAGGCAACAGTTCTGATTTATGATCAC 1047
QY 781 ATATTTGCCAGTAAAGCAGTGTGAAATGCGCAATTCAGCTTATCCTAAGACCTT 840
DB 1048 ATATTTGCCAGTAAAGCAGTGTGAAATGCGCAATTCAGCTTATCCTAAGACCTT 1107
QY 841 ATCAAAAGATCTTCATGATATCAAGCAGAAATTCCTGCGTAAATGCGATTTGTC 900
DB 1108 ATCAAAAGATCTTCATGATATCAAGCAGAAATTCCTGCGTAAATGCGATTTGTC 1167
QY 901 ACCCCATCTTTAGCATTCCTTTGCCCCCTCATTTTGAAGATCTGATGCTTCCACT 960
DB 1168 ACCCATCTTTAGCATTCCTTTGCCCCCTCATTTTGAAGATCTGATGCTTCCACT 1227
QY 961 CCAGTGTAAAGCTGCTGATGATGCTGATGATGATGATGATGATGATGATGATG 1020
DB 1228 CCAGTGTAAAGCTGCTGATGATGCTGATGATGATGATGATGATGATGATGATG 1287
QY 1021 GAAGATGTTGTGAAGATTAAGAGAGTTCGAAATATGAGCCAGTGTATCTCA 1080
DB 1288 GAAGATGTTGTGAAGATTAAGAGAGTTCGAAATATGAGCCAGTGTATCTCA 1347
QY 1081 CTTGTTCCAAAGAAATCTGCGCAGAGCAAGTCTTGTGATGATGCAATGCTGCT 1140
DB 1348 CTTGTTCCAAAGAAATCTGCGCAGAGCAAGTCTTGTGATGATGCAATGCTGCT 1407
QY 1408 GATTCGAAAGCTGCGCAGAAATTAAGTGAAGATGTTGATGGAAGTTGTGTG 1467
DB 1201 GCTACATTTTACCGCTGAGTGCCTTCAAGAGGGATATCTGTATCAAACTTGTCTTAA 1260
QY 1468 GCTACATTTTACCGCTGAGTGCCTTCAAGAGGGATATCTGTATCAAACTTGTCTTAA 1527
DB
RESULT 7
ADL45238
ID ADL45238 standard; DNA; 4065 BP.
XX
XX ADL45238;
XX
XX 20-MAY-2004 (first entry)
XX
XX Human ovarian cancer DNA marker #19128.
XX
XX Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.
XX
XX Homo sapiens.
XX
XX WO200170979-A2.
XX
XX 27-SEP-2001.
XX

PF	21-MAR-2001; 2001WO-US0009126.
XX	
PR	21-MAR-2000; 2000US-0191031P.
PR	25-MAY-2000; 2000US-0207124P.
PR	15-JUN-2000; 2000US-0211594P.
PR	07-JUL-2000; 2000US-0216820P.
PR	25-JUL-2000; 2000US-0220651P.
PR	21-DEC-2000; 2000US-0257672P.
XX	
PA	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX	
PI	Lee J, Little J;
XX	
DR	WPI; 2001-611502/70.
XX	
PT	Novel isolated nucleic acid molecules (markers) overexpressed in ovarian
PT	cancer cells as compared to their normal non-cancerous ovarian cells are
PT	used to characterize stage, grade, histological type of ovarian cancer.
XX	
XX	Disclosure; SEQ ID NO 19128; 106pp; English.

Db	448	ACCAACCGGGGCTGGGGCTCTGCGCGCGAGATATGGTTTCGCAAAAGAGGGGCGGGCTG	507
Qy	241	GAACGATTGCAGGGTCACAGAAACATGCTGACTTTGATATGAGAGTTTACATCCACTTT	300
Db	508	GAACGATTGCAGGGTCACAGAAACATGCTGACTTTGATATGAGAGTTTACATCCACTTT	567
Qy	301	TCTCCAAATGAGCCATCAAGCTGCTGTTGCTTAACTCCAGATATGTCAGTGTTCGGAA	360
Db	568	TCTCCAAATGAGCCATCAAGCTGCTGTTGCTTAACTCCAGATATGTCAGTGTTCGGAA	627
Qy	361	TTGCTCTATATATTCAGTCAACAGGAGTTGTTCCATATGATATGATACAGCAATGCGCCGA	420
Db	628	TTGCTCTATATATTCAGTCAACAGGAGTTGTTCCATATGATATGATACAGCAATGTCGCCGA	687
Qy	421	GATGTTTGAAGGCCCTTGCTTTTCTTCATCATGAGGGCATATGTCATGCGGACCTCAA	480
Db	688	GATGTTTGAAGGCCCTTGCTTTTCTTCATCATGAGGGCATATGTCATGCGGACCTCAA	747
Qy	481	CCAAGTACATATTTGTGAGATGTCAGAGAAATGAATGTTTAACTCATTTGACTTTGACTT	540
Db	748	CCAGTACATATTTGTGAGATGTCAGAGAAATGAATGTTTAACTCATTTGACTTTGACTT	807
Qy	541	AGCTTCAAAGAGGCAATCAGAGATGAATGATATTCAGACAGACGGGATTCGGGCTCCA	600
Db	808	AGCTTCAAAGAGGCAATCAGAGATGAATGATATTCAGACAGACGGGATTCGGGCTCCA	867
Qy	601	GAAGCAGATTCGCAAAATTCGCTTGCGCCAGGCTGCGCTGACAGATGATACAGAAATGATCC	660
Db	868	GAAGCAGATTCGCAAAATTCGCTTGCGCCAGGCTGCGCTGACAGATGATACAGAAATGATCC	927
Qy	661	TCAGTGTGATATCTGTGAGAGCTTAGGAATCATTTTATCTGAGAAATGTTCTCAGGAATGAAT	720
Db	928	TCAGTGTGATATCTGTGAGAGCTTAGGAATCATTTTATCTGAGAAATGTTCTCAGGAATGAAT	987
Qy	721	CTGAAACATAAGTCAGATCTCAGGAATGGAAGGCAACAGTTCTGCTATATTTGATATCAC	780
Db	988	CTGAAACATAAGTCAGATCTCAGGAATGGAAGGCAACAGTTCTGCTATATTTGATATCAC	1047
Qy	781	ATATTTGCGACGTAAAGCAGTGTGAATGCGCAATTCAGGCTATTCACCTTAAGAGACTT	840
Db	1048	ATATTTGCGACGTAAAGCAGTGTGAATGCGCAATTCAGGCTATTCACCTTAAGAGACTT	1107
Qy	841	ATCAAAAGCATGCTTCATGATGATCCAAAGCAGGAAGATTCCTGCTGAAATGGCATTTGTC	900
Db	1108	ATCAAAAGCATGCTTCATGATGATCCAAAGCAGGAAGATTCCTGCTGAAATGGCATTTGTC	1167
Qy	901	AGCCCATCTTAAAGATTCCTTTTGCCCTCATATTTGAAGATCTGATGATCTTCCCACT	960
Db	1168	AGCCCATCTTAAAGATTCCTTTTGCCCTCATATTTGAAGATCTGATGATCTTCCCACT	1227
Qy	961	CCAGTGTAAAGCTGCTGAATGTGCTGATGATGATTAATCTTGGGAATGAAGAGAAATAT	1020
Db	1228	CCAAGTGTAAAGCTGCTGAATGTGCTGATGATGATTAATCTTGGGAATGAAGAGAAATAT	1287
Qy	1021	GAAGATGTTGAGAGATGTAAGAGAGAGTGTCAAAATATATGACCAAGTGTATCTCTA	1080
Db	1288	GAAGATGTTGAGAGATGTAAGAGAGAGTGTCAAAATATATGACCAAGTGTATCTCTA	1347
Qy	1081	CTTGTTCCAAAGAAAAATCTTGCGCAGAGGAACAAGCTTTGTTGAGTATGCAAAATGCTGGT	1140
Db	1348	CTTGTTCCAAAGAAAAATCTTGCGCAGAGGAACAAGCTTTGTTGAGTATGCAAAATGCTGGT	1407
Qy	1141	GATTCCAAAGCTGCGCAGAAATTAATCTGAATGGAAGATGTTTGAATGGGAATTTGTTGTG	1200
Db	1408	GATTCCAAAGCTGCGCAGAAATTAATCTGAATGGAAGATGTTTGAATGGGAATTTGTTGTG	1467
Qy	1201	GCTACATTCTAACCGCTGAGTGCCTTACAGAGGGGATATCTGATCAAACTTGCTTTAA	1266
Db	1468	GCTACATTCTAACCGCTGAGTGCCTTACAGAGGGGATATCTGATCAAACTTGCTTTAA	1527

ID AAS06734 standard; cDNA; 1260 BP.
 XX
 AC AAS06734;
 XX
 DT 12-SEP-2001 (first entry)
 DE Polynucleotide sequence encoding human protein kinase #34.
 XX
 KM Human; protein kinase; PK; STK; cancer; cardiovascular disease;
 KM metabolic disorder; immune related disease; neurological disorder;
 KM neurodegenerative disorder; inflammatory disorder; infectious disease;
 KM reproductive disorder; gene therapy; ss.
 XX
 OS Homo sapiens.
 PN W0200138503-A2.
 PD 31-MAY-2001.
 PF 22-NOV-2000; 2000MO-US032085.
 PR 24-NOV-1999; 99US-0167482P.
 PA (SUGR-) SUGEN INC.
 PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
 PI Flanagan P, Clary D;
 DR WPI; 2001-343950/36.
 XX P-PSDB; AAU03534.
 PT Nucleic acids encoding human kinase polypeptides, useful for preventing
 PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
 PT neuronal-associated diseases, and microbial infections.
 XX
 PS Example 1; Fig 1; 433bp; English.
 CC AAS06701-AAS06757 encode for novel human protein kinases #1-57. The novel
 CC protein kinases have been identified as members of the tyrosine or
 CC serine/threonine kinase (PTK and STK) families. The polynucleotides
 CC encoding protein kinases and the polypeptides may be used in the
 CC prevention, diagnosis and treatment of diseases associated with
 CC inappropriate kinase expression. For example, they may be used to treat
 CC cancers (especially cancers of hematopoietic origin), cardiovascular
 CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
 CC immune related diseases (e.g. rheumatoid arthritis), neurological
 CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
 CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
 CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
 CC Additionally, polynucleotides encoding protein kinases may be used for
 CC gene therapy, and as DNA probes in diagnostic assays. The protein kinase
 CC polypeptides may be used as antigens in the production of antibodies
 CC against the protein kinases and in assays to identify modulators of
 CC protein kinase expression and activity
 XX
 SQ Sequence 1260 BP; 311 A; 279 C; 335 G; 335 T; 0 U; 0 Other;
 Query Match 99.5%; Score 1253.6; DB 4; Length 1260;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1256; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 181 ACCACCGGGGCTCGGCGCTCTGCGCCGAGATAGTGTTCGCGAAGAGAGGCGCGCTG 240
 DB 181 ACCACCGGGGCTCGGCGCTCTGCGCCGAGATAGTGTTCGCGAAGAGAGGCGCGCGCTG 240
 QY 241 GAACAGTTGACAGGCTCAAGAAACATCGTACTTTGTATGAGTGTTCACATTCACATTT 300
 DB 241 GAACAGTTGACAGGCTCAAGAAACATCGTACTTTGTATGAGTGTTCACATTCACATTT 300
 QY 301 TCTCCAAATGTCGATCAACGCTGTCTGTGTTGAATCTCGGATGTCAGTGTTCGGA 360
 DB 301 TCTCCAAATGTCGATCAACGCTGTCTGTGTTGAATCTCGGATGTCAGTGTTCGGA 360
 QY 361 TTGCTCTTATTTCCAGTCAACAGGAGTGTTCATGAGATGATACAGATTCGCCCGA 420
 DB 361 TTGCTCTTATTTCCAGTCAACAGGAGTGTTCATGAGATGATACAGATTCGTGCCGA 420
 QY 421 GATGTTTTGAGAGCCCTTGCTTTTCTTCATCATGAGGCTATGTCATGCGGACCTCAAA 480
 DB 421 GATGTTTTGAGAGCCCTTGCTTTTCTTCATCATGAGGCTATGTCATGCGGACCTCAAA 480
 QY 481 CCAAGTAACTATTTGTGAGTGCAGAGATGAAATGTTTAACTGATTCATTTGAGCTT 540
 DB 481 CCAAGTAACTATTTGTGAGTGCAGAGATGAAATGTTTAACTGATTCATTTGAGCTT 540
 QY 541 AGCTTCAAGAGAGCAATCAGGATGTAAAGTATATTCAGACAGACGAGTATCGGCTCCA 600
 DB 541 AGCTTCAAGAGAGCAATCAGGATGTAAAGTATATTCAGACAGACGAGTATCGGCTCCA 600
 QY 601 GAAGCAGAAATTGCAAAATTTGCTTGCCCGAGGCTGCGCAGAGTGAACAGATGTAC 660
 DB 601 GAAGCAGAAATTGCAAAATTTGCTTGCCCGAGGCTGCGCAGAGTGAACAGATGTAC 660
 QY 661 TCAGCTGTGATCTGTGAGAGCTTGAAGATATTTTAACTGGAATGTTCTCAGGAATGAA 720
 DB 661 TCAGCTGTGATCTGTGAGAGCTTGAAGATATTTTAACTGGAATGTTCTCAGGAATGAA 720
 QY 721 CTGAATCATACAGTCAAGATCTCAGAAATGGAAGCAACAGTTCTGTATTTATGATCAC 780
 DB 721 CTGAATCATACAGTCAAGATCTCAGAAATGGAAGCAACAGTTCTGTATTTATGATCAC 780
 QY 781 ATATTTGCGAGTAAAGCAATGATGATGCGCAATTCAGCTTATCATTAAAGACCTT 840
 DB 781 ATATTTGCGAGTAAAGCAATGATGATGCGCAATTCAGCTTATCATTAAAGACCTT 840
 QY 841 ATCAAAAGATGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
 DB 841 ATCAAAAGATGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
 QY 901 AGCCCATTTCTTGAAGATGCTTGTGCGCCCTCATATTTGAAGATCTGATGCTTCCACT 960
 DB 901 AGCCCATTTCTTGAAGATGCTTGTGCGCCCTCATATTTGAAGATCTGATGCTTCCACT 960
 QY 961 CCAAGTCTAAGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
 DB 961 CCAAGTCTAAGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
 QY 1021 GAAAGATGTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
 DB 1021 GAAAGATGTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
 QY 1081 CTTGTTCCAAAGAAATCTGCGACAGACAGATCTTGTGTGATGATGATGATGATGATG 1140
 DB 1081 CTTGTTCCAAAGAAATCTGCGACAGACAGATCTTGTGTGATGATGATGATGATGATG 1140
 QY 1141 GATTCGAAAGCTGCGACAGAAATCTGAGAGATGTTGATGAGAAATGTTGTTGTG 1200
 DB 1141 GATTCGAAAGCTGCGACAGAAATCTGAGAGATGTTGATGAGAAATGTTGTTGTG 1200
 QY 1201 GCTACATTTACCGCTGAGTGCCTCAAGAGGAGATCTGTATCAAACTTGTGCTTAA 1260
 DB 1201 GCTACATTTACCGCTGAGTGCCTCAAGAGGAGATCTGTATCAAACTTGTGCTTAA 1260

RESULT 9
 AAD18818
 ID AAD18818 standard; cDNA; 1824 BP.
 XX
 AC AAD18818;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human Kinase (PKIN) -3 CDNA.
 XX
 KW Human kinase; PKIN; gene therapy; adenocarcinoma; immune disorder; gout; cancer; allergy; sarcoma; leukemia; acquired immune deficiency syndrome; AIDS; Addison's disease; microbial infection; inflammation; osteoporosis; atherosclerosis; cardiovascular disease; myocardial infarction; anaemia; myasthenia gravis; cirrhosis; cataract; growth and development disorder; seizure disorder; pulmonary embolism; Gaucher's disease; lipid disorder; lipid storage disease; Pick's disease; Tay-Sachs disease; renal disease; asthma; obesity; restorative therapy; cytostatic; immunomodulatory; antimicrobial; cardiovascular; antiinflammatory; vaccine; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 169..1428
 FT /tag= a
 FT /product= "Human PKIN-3 protein"
 XX
 PN MO200181555-A2.
 XX
 XX 01-NOV-2001.
 XX
 XX 20-APR-2001; 2001WO-US012992.
 XX
 XX 20-APR-2000; 2000US-0199021P.
 XX 28-APR-2000; 2000US-0200226P.
 XX 05-MAY-2000; 2000US-0202339P.
 XX 11-MAY-2000; 2000US-0203505P.
 XX 18-MAY-2000; 2000US-0205564P.
 XX 26-MAY-2000; 2000US-0207739P.
 XX 01-JUN-2000; 2000US-0208795P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 XX Yue H, Gandhi AR, Tribouley CM, Kearney L, Griffin JA, Nguyen DB;
 XX Bandman O, Lu DM, Lal P, Burford N, Khan FA, Walla NK, Yao MG;
 XX Patterson C, Burdill JP, Marcus GA, Zingler RA, Reardon SA, Lu Y;
 XX Pollocky JL, Thornton M, Tang YT, Hefalia A, Elliott VS, Baughn MR;
 XX Walsh RT, Raskumar J, Borowsky ML, Au-Young J, Hillman JL;
 XX Gururajan R;
 XX
 XX MPI; 2001-611740/70.
 XX P-PSDB; AAE11769.
 DR
 DR Human kinases and nucleic acids, useful for preventing diagnosing and
 PT treating cancers, inflammation and immune disorders.
 PT
 XX
 XX Claim 5; Page 152; 166pp; English.
 PS
 XX The present invention relates to human kinases (PKIN) and the nucleic
 CC acids encoding them. PKIN is used as vaccine and in gene therapy. PKIN is
 CC used in the prevention, diagnosis and treatment of diseases cancers,
 CC adenocarcinoma, leukemia, sarcoma, immune disorder, Addison's disease,
 CC acquired immune deficiency syndrome (AIDS), anaemia, asthma, allergies,
 CC gout, microbial infections, cardiovascular disease and/or inflammation,
 CC myasthenia gravis, atherosclerosis, cirrhosis, osteoporosis, myocardial
 CC infection, cataract, growth and development disorder, seizure disorder,
 CC pulmonary embolism, Gaucher's disease, lipid disorder, lipid storage
 CC disease, Pick's disease, Tay-Sachs disease, renal disease and obesity.
 CC PKIN may be used to treat disorders associated with decreased PKIN
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of PKIN by expressing inactive proteins or to
 CC supplement the patient's own production of PKIN. PKIN nucleic acids may be
 CC used to produce the PKIN polypeptide, by inserting the nucleic acids into

CC a host cell and culturing the cell to express the protein. PKIN nucleic
 CC acid and its complementary sequences may also be used as DNA probes in
 CC diagnostic assays to detect and quantitate the presence of similar
 CC nucleic acid sequences in samples and therefore which patients may be in
 CC need of restorative therapy. The present sequence is human PKIN-3 CDNA
 XX
 XX SQ Sequence 1824 BP; 434 A; 416 C; 472 G; 502 T; 0 U; 0 Other;
 Query Match 99.5%; Score 1253.6; DB 4; Length 1824;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1256; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 ATGGCGGAGATCCGCTGCGCTTGGGCGCGAGACCGCGCGCTTTCTGAGAGCTTCGGG 60
 DB 169 ATGGCGGAGATCCGCTGCGCTTGGGCGCGAGACCGCGCGCTTTCTGAGAGCTTCGGG 228
 QY 61 CGGCTGTGCAGATTCAGAGCGGCTCTGAGGATGAGGCGCTCTCGGCTGATGAGGAT 120
 DB 229 CGGCTGTGCAGATTCAGAGCGGCTCTGAGGATGAGGCGCTCTCGGCTGATGAGGAT 288
 QY 121 CGCTGCTGCGGACCTTGGCTGCGCCCGCGCGCTTCAGGACATTCCTTGGCGGACAGA 180
 DB 289 CGCTGCTGCGGACCTTGGCTGCGCCCGCGCGCTTCAGGACATTCCTTGGCGGACAGA 348
 QY 181 ACCACCGGCGCTGCGGCTTGGCGCGAGATGATGCTTCGGAAGAAGAGGCGGCGCTG 240
 DB 349 ACCACCGGCGCTGCGGCTTGGCGCGAGATGATGCTTCGGAAGAAGAGGCGGCGCTG 408
 QY 241 GAACAGTTGCAGGATTCAGAGAACATCGGACTTTGATGAGGATTCATTCACCTT 300
 DB 409 GAACAGTTGCAGGATTCAGAGAACATCGGACTTTGATGAGGATTCATTCACCTT 468
 QY 301 TCTCCAAATGTGCATCAAGCTGTCTTGTGCTTGAATCTTGGATGTCAATGTTTCGGA 360
 DB 469 TCTCCAAATGTGCATCAAGCTGTCTTGTGCTTGAATCTTGGATGTCAATGTTTCGGA 528
 QY 361 TTGCTCTTATTTCCAGTACAGGAGTGTTCATGATGATGATACAGATTCGCGCCGA 420
 DB 529 TTGCTCTTATTTCCAGTACAGGAGTGTTCATGATGATGATACAGATTCGCGCCGA 588
 QY 421 GATGTTTGGAGGCGCTTCTTCTCATCATGAGGCTATGTCAAGCGGACCTCAAA 480
 DB 589 GATGTTTGGAGGCGCTTCTTCTCATCATGAGGCTATGTCAAGCGGACCTCAAA 648
 QY 481 CCACGTAACATATTTGGAGTGCAGAGATGATGATTTTAACTGATTCATTTGACTT 540
 DB 649 CCACGTAACATATTTGGAGTGCAGAGATGATGATTTTAACTGATTCATTTGACTT 708
 QY 541 ACCTTCAAGAAAGGCAATCAGATGTAAAGTATTTACAGACAGAGCGGATTCGGCTCA 600
 DB 709 ACCTTCAAGAAAGGCAATCAGATGTAAAGTATTTACAGACAGAGCGGATTCGGCTCA 768
 QY 601 GAAGCAGAAATTCGCAAAATTCCTTGGCCGAGCTGCGCAGAGTATACAGATGTACC 660
 DB 769 GAAGCAGAAATTCGCAAAATTCCTTGGCCGAGCTGCGCAGAGTATACAGATGTACC 828
 QY 661 TCAGCTGTGATCTGTGAGAGCTTACGAAATCATTTTACTGAAATGTTCTCAGAAATGAA 720
 DB 829 TCAGCTGTGATCTGTGAGAGCTTACGAAATCATTTTACTGAAATGTTCTCAGAAATGAA 888
 QY 721 CTGAAACATACATGATCTGAGATCTGAGAAATGGAAGCAACAGTTCTGCTATTTGATCAC 780
 DB 889 CTGAAACATACATGATCTGAGATCTGAGAAATGGAAGCAACAGTTCTGCTATTTGATCAC 948
 QY 781 AATTTTGCAGTAAGAGATGAGTGAATCCGGAATTCAGAGCTATCAGCTTAAGAGACCTT 840
 DB 949 AATTTTGCAGTAAGAGATGAGTGAATCCGGAATTCAGAGCTATCAGCTTAAGAGACCTT 1008
 QY 841 ATCAAAAGCATCTTCATGATGATCCAGACAGAAATTCCTGCTGAATGCGATTTGCG 900
 DB 1009 ATCAAAAGCATCTTCATGATGATCCAGACAGAAATTCCTGCTGAATGCGATTTGCG 1068
 QY 901 AGCCCATCTTTAGCATTCCTTTTGCCTCCTCATATTGAAGATCTGCTATGCTTCCACT 960

Db 1069 AGCCCATCTTACGATCTCTTGGCCCTCATATGGAAGATCGTCATGCTTCCACT 1128
Qy 961 CCAGTCTAAGCTCTGTAATGCTGATGATGATTAATCTGGGAATGAAGGAATAT 1020
Db 1129 CCAAGTCTAAGCTCTGTAATGCTGATGATGATTAATCTGGGAATGAAGGAATAT 1188
Qy 1021 GAAGATGTTGTGAAGATGTAAGAGAGATGCAAAAATATGACAGTGTATCTTA 1080
Db 1189 GAAGATGTTGTGAAGATGTAAGAGAGATGCAAAAATATGACAGTGTATCTTA 1248
Qy 1081 CTTGTTCCAAAGGAATCTCTGACAGAGCAAGTCTTGTGATGCAAAATGCTGT 1140
Db 1249 CTTGTTCCAAAGGAATCTCTGACAGAGCAAGTCTTGTGATGCAAAATGCTGT 1308
Qy 1141 GATTCCAAAGCTCTGACAGAGCAAGTCTTGTGATGCAAAATGCTGT 1200
Db 1309 GATTCCAAAGCTCTGACAGAGCAAGTCTTGTGATGCAAAATGCTGT 1368
Qy 1201 GCTACATTCACCCGCTGAGTGCCTCAAGAGGGGATATCTGTATCAAACTTTTAA 1260
Db 1369 GCTACATTCACCCGCTGAGTGCCTCAAGAGGGGATATCTGTATCAAACTTTTAA 1428

RESULT 10
ADQ24401
ID ADQ24401 standard; DNA; 3998 BP.
AC ADQ24401;
XX 26-AUG-2004 (first entry)
XX Human soft tissue sarcoma-upregulated DNA - SEQ ID 7221.
XX soft tissue sarcoma; cytosolic; gene therapy; vaccine; screening; human;
XX de.
XX Homo sapiens.
XX WO2004048938-A2.
XX 10-JUN-2004.
XX 26-NOV-2003; 2003WO-US038193.
XX 26-NOV-2002; 2002US-0429739P.
XX (PROT-) PROTEIN DESIGN LABS INC.
XX Aziz N, Ginsburg WM, Zlotnik A;
XX WPI; 2004-441208/41.
XX
XX Early detection of soft tissue sarcoma comprises determining expression
XX of a gene in a first soft tissue sample and a normal soft tissue sample
XX and comparing the gene expression, also useful in treating soft tissue
XX sarcoma.
XX
XX Example 2; SEQ ID NO 7221; 210pp; English.

The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytosolic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC DNA of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.

XX
SQ Sequence 3998 BP; 1066 A; 777 C; 885 G; 1255 T; 0 U; 15 Other;
Query Match 93.1%; Score 1173.4; DB 12; Length 3998;
Best Local Similarity 96.1%; Pred. No. 0;
Matches 1214; Conservative 0; Mismatches 46; Indels 3; Gaps 1;
Qy 1 ATGGCGGATCCGGCTGCGCCCTGCGAGCGCGCGCTTCTGAGGCTTCCGG 60
Db 160 ATGGCGGATCCGGCTGCGCCCTGCGAGCGCGCGCTTCTGAGGCTTCCGG 219
Qy 61 CGGCTGTGGCAGTACAGAGCCGTCTGGAGCGGCTCTCCGCTCGGTATCGGTT 120
Db 220 CGGCTGTGGCAGTACAGAGCCGTCTGGAGCGGCTCTCCGCTCGGTATCGGTT 279
Qy 121 CGGCTGTGGCAGTACAGAGCCGTCTGGAGCGGCTCTCCGCTCGGTATCGGTT 180
Db 280 CGGCTGTGGCAGTACAGAGCCGTCTGGAGCGGCTCTCCGCTCGGTATCGGTT 339
Qy 181 ACCACCGGGGCTCGGGCTCTGCGCCGAGTATGTTCCGCAAGAGAGGCGCGCTG 240
Db 340 ACCACCGGGGCTCGGGCTCTGCGCCGAGTATGTTCCGCAAGAGAGGCGCGCTG 399
Qy 241 GAACAGTTGCAAGGTTCAGAGAAACATCGTACTTTGATGAGGTTTTCAATCCATT 300
Db 400 AAGTCTGTGTGACCTTAACCGAGAGTGACTTTGTATGAGGTTTTCAATCCAGGTT 459
Qy 301 TCT---CCAAATGTGCCATCAGCTGTCTGTGCTTGAATCTCTGAGTCACTTTG 357
Db 460 GTTCTCCAAAGTGTCCATCAGCTGTCTGTGCTTGAATCTCTGAGTCACTTTG 519
Qy 358 GAATGCTCTTATATTCAGTCAAGAGGTTGTCATGATGATGATCAGATTCGCGC 417
Db 520 GAATGCTCTTATATTCAGTCAAGAGGTTGTCATGATGATGATCAGATTCGCGC 579
Qy 418 CGAGATGTTTGAAGCCCTTGTCTTCTCATCATGAGGGCTATGTCATGCGGACCTC 477
Db 580 CGAGATGTTTGAAGCCCTTGTCTTCTCATCATGAGGGCTATGTCATGCGGACCTC 639
Qy 478 AAACCACTGATATTTGTGAGTGCAGAGATGAATGTTTAACTCATGACTTTGGA 537
Db 640 AAACCACTGATATTTGTGAGTGCAGAGATGAATGTTTAACTCATGACTTTGGA 699
Qy 538 CTTAGCTTCAAGAGGCAATCAGATGTAAGTATATTCAGACAGAGGATATGGGCT 597
Db 700 CTTAGCTTCAAGAGGCAATCAGATGTAAGTATATTCAGACAGAGGATATGGGCT 759
Qy 598 CCAGAAAGCAATTCGCAAAATGCTTGGCCCGAGGCTGCGCAGATGATCAGAAATG 657
Db 760 CCAGAAAGCAATTCGCAAAATGCTTGGCCCGAGGCTGCGCAGATGATCAGAAATG 819
Qy 658 ACTCAGCTGTGATCTGTGAGGCTTGGAGATCATTTATCTGGAATGTTCTCAGAAATG 717
Db 820 ACTCAGCTGTGATCTGTGAGGCTTGGAGATCATTTATCTGGAATGTTCTCAGAAATG 879
Qy 718 AAATGAAACATACATCAGATCTCAGGAATGGAAGGCAACAGTTCTGTATTTGAT 777
Db 880 AAATGAAACATACATCAGATCTCAGGAATGGAAGGCAACAGTTCTGTATTTGAT 939
Qy 778 CACATATTTTCCAGTAAAGCAGTGTGATGCGCAATTCAGCTTACCTTAAGAGAC 837
Db 940 CACATATTTTCCAGTAAAGCAGTGTGATGCGCAATTCAGCTTACCTTAAGAGAC 999
Qy 838 CTTATCAAAAGATCTTATGATATATCCAGCAGAGATTCCTGCTGAATGCAATG 897
Db 1000 CTTATCAAAAGATCTTATGATATATCCAGCAGAGATTCCTGCTGAATGCAATG 1059
Qy 898 TCGAGCCATTTTATGATCTTTTGGCCCTCATTTGAAGATCTGTATGCTTCC 957
Db 1060 TCGAGCCATTTTATGATCTTTTGGCCCTCATTTGAAGATCTGTATGCTTCC 1119
Qy 958 ACTCAGTCTAAGCTGCTGATGCTGATGATGATTTCTTGGAAATGAAGAGAA 1017

DB 1120 ACTCAGTCTAGACTGCTGTAATGTCTGATGATGATTACTTGAGAAATGAGAGAA 1179
QY 1018 TATGAAGATGTTTATGAAAGATGTAAAGAGAGAGTCAAAAATATGACCAGTGTATCT 1077
DB 1180 TATGAAGATGTTTATGAAAGATGTAAAGAGAGAGTCAAAAATATGACCAGTGTATCT 1239
QY 1078 CTACTGTTTCCAAAGAGAAATCTCTGACAGAGCAAGTCTTTTGTATGATGCAAAATGCT 1137
DB 1240 CTACTGTTTCCAAAGAGAAATCTCTGACAGAGCAAGTCTTTTGTATGATGCAAAATGCT 1299
QY 1138 GGATGATTCACAAAGCTGCGAGAAATTAATGACTGGAAGATGTTTATGAGGAAGTTGTT 1197
DB 1300 GGATGATTCACAAAGCTGCGAGAAATTAATGACTGGAAGATGTTTATGAGGAAGTTGTT 1359
QY 1198 GTGGCTACATTTTACCCGCTGAGTGCCTACAAAGAGGAGATATCTGTATCAAACTTGCTT 1257
DB 1360 GTGGCTACATTTTACCCGCTGAGTGCCTACAAAGAGGAGATATCTGTATCAAACTTGCTT 1419
QY 1258 TAA 1260
DB 1420 TAA 1422

RESULT 11
ADB52857
ID ADB52857 standard; DNA: 3244 BP.
XX
AC ADB52857;
XX
DT 04-DEC-2003 (first entry)
XX
DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3399.
XX
KM toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KM toxicity marker; toxicity progression; drug screening;
KM primary rat hepatocyte toxicity modelling; gene; ds.
XX
OS Rattus norvegicus.
XX
PN MO2003065993-A2.
XX
PD 14-AUG-2003.
XX
PF 04-FEB-2003; 2003MO-US003482.
XX
PR 04-FEB-2002; 2002US-0353171P.
PR 13-MAR-2002; 2002US-0363534P.
PR 08-APR-2002; 2002US-0370248P.
PR 10-APR-2002; 2002US-0371134P.
PR 10-APR-2002; 2002US-0371135P.
PR 10-APR-2002; 2002US-0371150P.
PR 11-APR-2002; 2002US-0371433P.
PR 19-APR-2002; 2002US-0373601P.
PR 19-APR-2002; 2002US-0373602P.
PR 22-APR-2002; 2002US-0374139P.
PR 08-MAY-2002; 2002US-0378370P.
PR 09-MAY-2002; 2002US-0378652P.
PR 09-MAY-2002; 2002US-0378653P.
PR 09-MAY-2002; 2002US-0378655P.
PR 09-JUL-2002; 2002US-0394230P.
PR 09-JUL-2002; 2002US-0394253P.
PR 04-SEP-2002; 2002US-0407688P.
PR 28-JAN-2003; 2003US-0442900P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Mendick D, Porter M, Johnson K, Higgs B, Caastle A, Orr M;
PI Blashoff M;
XX
DR WPI; 2003-731472/69.
XX
PT Determining if a compound induces a toxic effect on a tissue or cell, for
PT identifying hepatotoxic compounds, comprises comparing a gene expression

PT profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT mean values.
XX
XX
PS Claim 44; SEQ ID NO 3399; 874bp; English.
XX
CC The present invention describes a method for determining whether a
CC compound induces a toxic effect on a tissue or cell. The method comprises
CC preparing a gene expression profile of a tissue or cell sample exposed to
CC the compound, and comparing the gene expression profile to a database
CC comprising data or information on the Tox mean and non-Tox mean value.
CC The method is useful for predicting or identifying at least one toxic
CC effect, particularly hepatotoxicity, of a test or unknown compound. The
CC genes listed in the specification are useful as diagnostic or toxicity
CC markers for the prediction or identification of the physiological state
CC of tissue or cell sample that has been exposed to a compound, or to
CC identify or predict the toxic effects of a compound or an agent. These
CC may also be used as markers for monitoring toxicity progression or for
CC drug screening. The present sequence represents a primary rat hepatocyte
CC toxicity modelling related gene sequence from the present invention.
XX
SQ Sequence 3244 BP; 800 A; 710 C; 819 G; 915 T; 0 U; 0 Other;
Query Match 87.8%; Score 1106.4; DB 10; Length 3244;
Best Local Similarity 92.4%; Pred. No. 0;
Matches 1164; Conservative 0; Mismatches 96; Indels 0; Gaps 0;
QY 1 ATGGCGGAGATCCGCGCTGCGCCTGCGGCGCGAGCCGCGCTTTTCTGAGAGCCTTCGCG 60
DB 162 ATGGCCGGGTCCGGCTCGCGCGCGGTGGGGCGCGAGCCGCGCTTCGAGAGCCTTCGCG 221
QY 61 CGGCTGTGCGAGTACAGAGCCGCTCTGGGTACCGGCTCTCCGCTCGGTATATGGGTT 120
DB 222 CGGCTGTGCGAGTACAGAGCCGCTACAGACCGGCTCTCCGCTCGGTATATGGGTTG 281
QY 121 CGGCTGTGCGAGACCTTGCGCTCGCCCCCGCGGCTCTACAGAGTTCCTTCGCGCAGAA 180
DB 282 CGGCTGTGCGAGACCTTCAGAGCTCGCCCCCGCGGCTCTACAGAGTTCCTTCGCGCAG 341
QY 181 ACCACCGGGGCTGCGGCTCTGCGCCGAGATGATTTCCGAAAGAGAGGGCGGCTG 240
DB 342 ACCACCGGGGCTGCGGCTCTGCGGCTGCGGATGATTTCCGAAAGAGAGGGCGGCTG 401
QY 241 GAACAGTTGCGAGGTACAGAAACATCGTGAATTTGTATGAGATTTTACATCCACTTT 300
DB 402 GAGCAGTTGCGAGGTACAGAAACATCGTGAATTTGTATGAGATTTTACATCCACTTC 461
QY 301 TCTCCAAATGTCGATCAAGCGCTGCTGTTGCTTGAATCTTGATATGTCGATTTCCGAA 360
DB 462 TCTCCAAATGTCGATCAAGCGCTGCTGTTGCTTGAATCTTGATATGTCGATTTCCGAA 521
QY 361 TTGCTCTTATATTCAGATCAACGAGGTGTTTCATGTGATGATACAGCAATTCGCCGA 420
DB 522 TTGCTCTTATATTCAGATCAACGAGGTGTTTCATGTGATGATACAGCAATTCGCCGA 581
QY 421 GATGTTTGGAGCCCTTGCTTTTCTTATCATATGAGGCTATGTCCATGCGGACCTTCAA 480
DB 582 GATGTTTGGAGCCCTTGCTTTTCTTATCATATGAGGCTATGTCCATGCGGACCTTCAA 641
QY 481 CCACGTATACATTTTGGAGTGCAGAGATGAATGTTTAACTGATGACTTTGGACTT 540
DB 642 CCACGTATACATCTGTGAGTGCAGAGATGAATGTTTAACTGATGACTTTGGACTT 701
QY 541 ACCTTCAAAGAGGCAATCAGATGTAAAGTATATTCAGACAGACGGGTATCGGCTCCA 600
DB 702 ACCTTCAAAGAGGCAATCAGATGTAAAGTATATTCAGACAGACGGGTATCGGCTCCT 761
QY 601 GAAGCAGAAATTCGAAATTCCTTGGCCCAAGGCTGCGCTGCAAGTATACAGAAATGACC 660
DB 762 GAAGCAGAAATTCGAAATTCCTTGGCCCAAGGCTGCGCTGCAAGTATACAGAAATGACC 821
QY 661 TCAGCTGTGATCTGTGAGGCTTAGAATCATTTTACTGGAATGTTTCAGAAATGAAA 720
DB 822 TCAGCTGTGATCTGTGAGGCTTAGAATCATTTTACTGGAATGTTTCAGAAATGAAA 881

```
QY 721 CTGAAACATACAGTCAGATCTCAGAAATGAGAAAGCAACAGTCTGCTATTATTGATCAC 780
DB 882 CTGAAACATACAGTCAGATCTCAGAAATGAGAAAGCAACAGTCTGCTATTATTGATCAC 941
QY 781 AATTTTCCAGTAAAGCAGTGTGTAATGCCGCAATTCACGCTTATCATCCTAAGACACTT 840
DB 942 AATTTTCCAGTAAAGCAGTGTGTAATGCCGCAATTCACGCTTATCATCCTAAGACACTT 1001
QY 841 AATCAAAAGCATGCTTATGATGATGATCAAGCAAGAAATCCCTGCTAAGATGATGCTG 900
DB 1002 AATCAAAAGCATGCTTATGATGATGATCAAGCAAGAAATCCCTGCTAAGATGATGCTG 1061
QY 901 AGCCCATTCCTTTAGCATTCCTTTGCCCTCATATTGAAGATGCTGCTCCACT 960
DB 1062 AGCCCATTCCTTTAGCATTCCTTTGCCCTCATATTGAAGATGCTGCTCCACT 1121
QY 961 CCAGTCTAAGACTGCTGTAATGTGTGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 1122 CCAGTCTAAGACTGCTGTAATGTGTGATGATGATGATGATGATGATGATGATGATGAT 1181
QY 1021 GAAGATGTTGTAAGATGTAAGATGTAAGATGTAAGATGTAAGATGTAAGATGTAAGAT 1080
DB 1182 GAAGATGTTGTAAGATGTAAGATGTAAGATGTAAGATGTAAGATGTAAGATGTAAGAT 1241
QY 1081 CTTGTTCCAAAGGAAATCCCTGCAAGAGCAAGTCTTTGATGATGATGATGATGATGAT 1140
DB 1242 CTTGTTCCAAAGGAAATCCCTGCAAGAGCAAGTCTTTGATGATGATGATGATGATGAT 1301
QY 1141 GATTCCAAAGCTGCGCAGAAATTAATGATGATGATGATGATGATGATGATGATGATG 1200
DB 1302 GATTCCAAAGCTGCGCAGAAATTAATGATGATGATGATGATGATGATGATGATGATG 1361
QY 1201 GCTACATTTTACCCGCTGAGTGTCTACAGAGGGGATTTCTGATCAAACTTGTCTTAA 1260
DB 1362 GCTACATTTTACCCGCTGAGTGTCTACAGAGGGGATTTCTGATCAAACTTGTCTTAA 1421
```

RESULT 12

AAH70764
ID AAH70764 standard; cDNA, 529 BP.

XX AAH70764;

XX 19-SEP-2001 (first entry)

DE Human cervical cancer marker nucleic acid 2038.

XX Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

OS Homo sapiens.

XX WO200142467-A2.

PD 14-JUN-2001.

PF 08-DEC-2000; 2000WO-US033312.

XX 08-DEC-1999; 99US-0169681P.

XX 21-DEC-1999; 99US-0171350P.

PR 14-MAR-2000; 2000US-0189315P.

PR 12-MAY-2000; 2000US-0203791P.

PR 09-JUN-2000; 2000US-0210600P.

PR 21-JUL-2000; 2000US-0220114P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Deede J, Berger A, Zhao X;

XX WPI; 2001-375006/39.

PT New isolated nucleic acid for diagnosing and treating cervical cancer and
for assessing and detecting compounds for treating the cancer.

XX Claim 1; Page 434; 1051pp; English.
PS The invention relates to novel genes (AAH68727-AAH73383) associated with
XX cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy

XX Sequence 529 BP; 156 A; 102 C; 132 G; 139 T; 0 U; 0 Other;

Query Match 36.4%; Score 458.6; DB 4; Length 529;

Best Local Similarity 97.8%; Pred. No. 7.1e-126;

Matches 486; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

QY 654 ATGTACTCAGCTGTTGATCTGTGAGCCTAGAGATCAATTTTACGAAATGTTCTCAGG 713

DB 22 AGTACCTCAGCTGTTGATCTGTGAGCCTAGAGATCAATTTTACGAAATGTTCTCAGG 81

QY 714 AATGAACTGAAACATPACAGTCAGATCTCAGGAATGAAAGCAACAGTTCTGATTAAT 773

DB 82 AATGAACTGAAACATPACAGTCAGATCTCAGGAATGAAAGCAACAGTTCTGATTAAT 141

QY 774 TGATCAGATTTTGGCAGTAAAGCAGTGTGAATGCCCAATTCAGCCTATCAGCTTAAG 833

DB 142 TGATCAGATTTTGGCAGTAAAGCAGTGTGAATGCCCAATTCAGCCTATCAGCTTAAG 201

QY 834 AGACCTTATCAAAAGCATGCTCATGATGATCAAGAGAAATTCGCTGAATGGC 893

DB 202 AGACCTTATCAAAAGCATGCTCATGATGATCAAGAGAAATTCGCTGAATGGC 261

QY 894 ATGTGACAGCCCATTCCTTACGATTCCTTTGCCCCCTCATATTGAAGATCTGTATGCT 953

DB 262 ATGTGACAGCCCATTCCTTACGATTCCTTTGCCCCCTCATATTGAAGATCTGTATGCT 321

QY 954 TCCCACTCAGTGTAAAGATGCTGAATGTGTGTGATGATGATTTATCTTGGAAATGAAGA 1013

DB 322 TCCCACTCAGTGTAAAGATGCTGAATGTGTGTGATGATGATTTATCTTGGAAATGAAGA 381

QY 1014 GGATATGAGA- TGTTGTAAGATGTAAGA- GAGTGCAGAAATATGAGCAAGTG 1071

DB 382 GGATATGAGA- TGTTGTAAGATGTAAGA- GAGTGCAGAAATATGAGCAAGTG 441

QY 1072 GTATCTTACTTGTTCAGAAAGAAATCTGTGCAAGAGCAAGTCTTGTGATGATGCA 1131

DB 442 GTATCTTACTTGTTCAGAAAGAAATCTGTGCAAGAGCAAGTCTTGTGATGATGCA 501

QY 1132 AATGCTGATTCGAA 1148

DB 502 AAGGCTGGGATTCAAA 518

RESULT 13

AAA02535
ID AAA02535 standard; cDNA, 722 BP.

XX AAA02535;

DT 19-MAY-2000 (first entry)

XX Human colon cancer cell line polynucleotide sequence SEQ ID NO:2526.

XX Human; colon cancer; tumour; diagnosis; gene expression product; probe;

XX detection; cancerous state; metastasis; identification; breast cancer;

XX oestrogen receptor-negative breast cancer; lung cancer; ss.

XX Homo sapiens.

XX WO9958675-A2.

XX 18-NOV-1999.
 PD 13-MAY-1999; 99WO-US010602.
 XX 14-MAY-1998; 98US-0085426P.
 PR 15-MAY-1998; 98US-0085537P.
 PR 15-MAY-1998; 98US-0085696P.
 PR 21-OCT-1998; 98US-0105234P.
 PR 27-OCT-1998; 98US-0105877P.
 XX (CHIR) CHIRON CORP.
 PA (HYSE-) HYSEQ INC.
 PI Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
 PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
 PI Lamsan G, Drmanac R, Ckvenjakov R, Dickson M, Drmanac S, Labat I;
 PI Leshkowitz D, Kita D, Garcia V, Jones LW, Strache-Crain B;
 DR WPI, 2000-126369/11.
 XX Polynucleotide library used to determine cancerous states of mammalian
 PT cells.
 XX Claim 1; Page 1020; 1097pp; English.

XX AAA00010 to AAA02716 represent polynucleotides isolated from cDNA
 CC libraries constructed from human colon cancer cell lines. The present
 CC invention also describes a method of detecting differentially expressed
 CC genes correlated with a cancerous state of a mammalian cell, comprising
 CC detecting at least one differentially expressed gene product in a test
 CC sample derived from a cell suspected of being cancerous, where detection
 CC of the differentially expressed gene product is correlated with a
 CC cancerous state of the cell from which the test sample was derived. The
 CC polynucleotide sequences can be used in a method for detecting
 CC differentially expressed genes correlated with a cancerous state of a
 CC mammalian cell. The polynucleotides can also be used as probes for
 CC detecting and mapping related genes. They can be used in diagnosis and
 CC prognosis of diseases and disorders (e.g. identification of pre-
 CC metastatic or metastatic cancerous states, stages of cancer, or
 CC responsiveness of cancer to therapy). This is particularly for breast
 CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
 CC negative breast cancer, lung cancer, and colon cancer

SQ Sequence 722 BP; 215 A; 129 C; 164 G; 193 T; 0 U; 21 Other;

Query Match 35.4%; Score 445.8; DB 3; Length 722;

Best Local Similarity 93.0%; Pred. No. 5.6e-122; Indels 15; Gaps 11;

Matches 581; Conservative 0; Mismatches 29; Indels 15; Gaps 11;

QY 454 GAGGGCTATGTCCATCGGACCTCAACACGTAATATTGTGAGTCAGAGAAATGAA 513
 DB 1 GAGGGCTATGTCCATCGGACCTCAACACGTAATATTGTGAGTCAGAGAAATGAA 60
 QY 514 TGTTTTAACTCAATTGACTTTGACTTGAAGAGGCAATGTAAGTAT 573
 DB 61 TGTTTTAACTCAATTGACTTTGACTTGAAGAGGCAATGTAAGTAT 120
 QY 574 ATTCAGACGAGGGGTATGGGGTCCAGAGGAAATTTGCAAAATGCTGGCCAGGCT 633
 DB 121 ATTCAGACGAGGGGTATGGGGTCCAGAGGAAATTTGCAAAATGCTGGCCAGGCT 179
 QY 634 GGCCTCAGAGTATACAGAAATGACTTGTGATCTGTGAGCCTTGAAGATCAT 693
 DB 180 GGCCTCAGAGTATACAGAAATGACTTGTGATCTGTGAGCCTTGAAGATCAT 239
 QY 694 TTAATGAAATGTTCTCAGAAATGAAATGAAATGAAATGAAATGAAATGAA 753
 DB 240 TTAATGAAATGTTCTCAGAAATGAAATGAAATGAAATGAAATGAAATGAA 299
 QY 754 GCAAGAG-TTTGGCTATTA-TTGATCACAATTTGGCAATTA--GCGATGGTGAATG 809
 DB 300 GCAAGAGTTTGGCTATTA-TTGATCACAATTTGGCAATTA--GCGATGGTGAATG 359

QY 810 CGCAATTCAGGCTATCACTTAAGACCTTATCAAAAGCATGCTTCAGATGATCC--A 867
 DB 360 CGCAATTCAGGCTATCACTTAAGACCTTATCAAAAGCATGCTTCAGATGATCCAA 419
 QY 868 ACAGAGAAATTCCTGCTGAATGCAAT-GTGACGCCATCTTTAGC-ATTCCTTTG 925
 DB 420 GCAGAGAAATTCCTGCTGAATGCAATGTCATGTCATCTTTAGC-ATTCCTTTG 479
 QY 926 CCCCTCATTTTGAAGATCTGCTGATGCTTCCACTCCAGT-GCTAAGACTGCTGAATG 984
 DB 480 CCCCTCATTTTGAAGATCTGCTGATGCTTCCACTCCAGT-GCTAAGACTGCTGAATG 539
 QY 985 CT-GGATGATGATTTCTTTGGGAATGAA--GAGGAATGGAAGATGTTGAAGATG 1040
 DB 540 CTGGGATGATGATTTCTTTGGGAATGAA--GAGGAATGGAAGATGTTGAAGATG 599
 QY 1041 -AAAAGAGAGTGTCAAAATATG 1064
 DB 600 AAAAGAGAGTGTGCAAAATATTTG 624

RESULT 14

AAH71966
 ID AAH71966 standard; cDNA; 470 BP.

AAH71966;

19-SEP-2001 (first entry)

DE Human cervical cancer marker nucleic acid 3240.

KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

XX Homo sapiens.

PN WO200142467-A2.

PD 14-JUN-2001.

XX 08-DEC-2000; 2000WO-US033312.

PF 08-DEC-1999; 99US-0168681P.

PR 21-DEC-1999; 99US-0171350P.

PR 14-MAR-2000; 2000US-0189315P.

PR 09-JUN-2000; 2000US-0203791P.

PR 21-JUL-2000; 2000US-0220114P.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Deede J, Berger A, Zhao X;

DR WPI, 2001-375006/39.

PT New isolated nucleic acid for diagnosing and treating cervical cancer and
 for assessing and detecting compounds for treating the cancer.

PS Claim 1; Page 636; 1051pp; English.

XX The invention relates to novel genes (AAH68727-AAH73383) associated with
 CC cervical cancer with cytostatic activity. The nucleic acids and encoded
 CC polypeptides are useful: to assess if a patient is afflicted with
 CC cervical cancer or has a pre-malignant condition; to monitor the
 CC progression of cervical cancer or a premalignant condition in a patient;
 CC and to select and/or assess the efficacy of a compound or therapy for
 CC inhibiting cervical cancer in a patient. The nucleic acids may also be
 CC useful for gene therapy

SQ Sequence 470 BP; 140 A; 91 C; 114 G; 125 T; 0 U; 0 Other;

Query Match 33.9%; Score 426.6; DB 4; Length 470;

Best Local Similarity 98.7%; Pred. No. 2.4e-116;

Matches 451; Conservative 0; Mismatches 4; Indels 2; Gaps 2;
QY 654 ATGTACCTCAGCTGTTGATCTGTGAGCCAGAAATCTTTCTGGAATGTTCCAGG 713
DB 14 AGGTACCTCAGCTGTTGATCTGTGAGCCAGAAATCTTTCTGGAATGTTCCAGG 73
QY 714 AATGAACCTGAACATATCAGTCTGAGATCTCAGAAATGGAAGCAACAGTTCTGCTATTAT 773
DB 74 AATGAACCTGAACATATCAGTCTGAGATCTCAGAAATGGAAGCAACAGTTCTGCTATTAT 133
QY 774 TGATCAATATTTGCCAGTAAGAGTGTGAATGCCGAATTCAGCCTATCACTTAAG 833
DB 134 TGATCAATATTTGCCAGTAAGAGTGTGAATGCCGAATTCAGCCTATCACTTAAG 193
QY 834 AGACCTTATGAAGAGTGTGATCTGATCCAAAGAGAAATTCCTGCTGAATGCG 893
DB 194 AGACCTTATGAAGAGTGTGATCTGATCCAAAGAGAAATTCCTGCTGAATGCG 253
QY 894 ATGTGACAGCCATTCCTTTAGCATTCCTTTGCCCTCATATTGAAGATCTGCTCATGCT 953
DB 254 ATGTGACAGCCATTCCTTTAGCATTCCTTTGCCCTCATATTGAAGATCTGCTCATGCT 313
QY 954 TCCCACTCCAGTGTCTAAGACTGCTGAATGTCTGATGATGATTATCTTGGAATGAAGA 1013
DB 314 TCCCACTCCAGTGTCTAAGACTGCTGAATGTCTGATGATGATTATCTTGGAATGAAGA 373
QY 1014 GGAATATGAAGATGTTGTGAAGATGTAAGAAGAGTGTCAAAAATATGACCAAGG 1071
DB 374 GGAATATGAAGATGTTGTGAAGATGTAAGAAGAGTGTCAAAAATATGACCAAGG 433
QY 1072 GTATCTCTACTGTGTTCCAAAGAAATCTGCGCAGAG 1108
DB 434 GTATCTCTACTGTGTTCCAAAGAAATCTGCGCAGAG 470
RESULT 15
AAH69101
ID AAH69101 standard; cDNA; 461 BP.
XX
AC AAH69101;
XX
DT 19-SBP-2001 (first entry)
XX
DE Human cervical cancer marker nucleic acid 375.
XX
KM Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN MO200142467-A2.
PD 14-JUN-2001.
XX
PF 08-DEC-2000; 2000MO-US033312.
XX
PR 08-DEC-1999; 99US-0169681P.
PR 21-DEC-1999; 99US-0171350P.
PR 14-MAR-2000; 2000US-0189315P.
PR 12-MAY-2000; 2000US-0203791P.
PR 09-JUN-2000; 2000US-0210600P.
PR 21-JUL-2000; 2000US-0220114P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PI Schlegel R, Deeds J, Berger A, Zhao X;
XX
XX WPI; 2001-375006/39.
XX
XX New isolated nucleic acid for diagnosing and treating cervical cancer and
PT for assessing and detecting compounds for treating the cancer.
XX
PS Claim 1; Page 168-169; 1051pp; English.
XX

CC The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy
XX
SQ Sequence 461 BP; 134 A; 91 C; 109 G; 127 T; 0 U; 0 Other;
Query Match 29.8%; Score 375.4; DB 4; Length 461;
Best Local Similarity 96.2%; Pred. No. 4.4e-101;
Matches 428; Conservative 0; Mismatches 11; Indels 6; Gaps 4;
QY 654 ATGTACCTCAGCTGTTGATCTGTGAGCCAGAAATCTTTCTGGAATGTTCCAGG 713
DB 18 AGGTACCTCAGCTGTTGATCTGTGAGCCAGAAATCTTTCTGGAATGTTCCAGG 77
QY 714 AATGAACCTGAACATATCAGTCTGAGATCTCAGAAATGGAAGCAACAGTTCTGCTATTAT 773
DB 78 AATGAACCTGAACATATCAGTCTGAGATCTCAGAAATGGAAGCAACAGTTCTGCTATTAT 137
QY 774 TGATCAATATTTGCCAGTAAGAGTGTGAATGCCGAATTCAGCCTATCACTTAAG 833
DB 138 TGATCAATATTTGCCAGTAAGAGTGTGAATGCCGAATTCAGCCTATCACTTAAG 197
QY 834 AGACCTTATGAAGAGTGTGATCTGATCCAAAGAGAAATTCCTGCTGAATGCG 893
DB 198 AGACCTTATGAAGAGTGTGATCTGATCCAAAGAGAAATTCCTGCTGAATGCG 257
QY 894 ATGTGACAGCCATTCCTTTAGCATTCCTTTGCCCTCATATTGAAGATCTGCTCATGCT 953
DB 258 ATGTGACAGCCATTCCTTTAGCATTCCTTTGCCCTCATATTGAAGATCTGCTCATGCT 316
QY 954 TCCCACTCCAGTGTCTAAGACTGCTGAATGTCTGATGATGATTATCTTGGAATGAAGA 1013
DB 317 TCCCACTCCAGTGTCTAAGACTGCTGAATGTCTGATGATGATTATCTTGGAATGAAGA 376
QY 1014 GGAATATGAAGATGTTGTGAAGATGTAAGAAGAGTGTCAAAAATATGACCAAG 1068
DB 377 GGAATATGAAGATGTTGTGAAGATGTAAGAAGAGTGTCAAAAATATGACCAAG 436
QY 1069 GTGTATCTCTACTGTGTTCCAAAG 1093
DB 437 GTGTATCTCTACTGTGTTCCAAAG 461

Search completed: November 29, 2004, 11:45:27
Job time : 656.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: November 29, 2004, 11:25:02 ; Search time 3159.5 Seconds
(without alignments)
4832.489 Million cell updates/sec

Title: US-10-798-532-4
Perfect score: 2206
Sequence: 1 MAGSGCAGMABPPRFLEAFG.....VATFYPPLSAVKGIVYQTLL 419

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues
Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_p2n.model -DRV=x1p
-O=/cgn2_1/USPTO.epool_p/US10798532/runat_29112004_112453_15901/app_query.faeta_1.1166
-DB=EST-QFMT=fastap -SUFFIX=rest -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITs=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10798532.@CEN_1_1_6628@runat_29112004_112453_15901 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -KONGLOG
-NO_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:.*
1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gsa1.*
9: gb_gsa2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2171	98.4	1763	3	AK013347 Mus muscu
2	1711	77.6	2846	3	AK030152 Mus muscu
3	1521.5	69.0	1079	4	BM468107 AGENCOURT
4	1271.5	57.6	1015	6	BY712344 BY712344
5	1242	56.3	917	4	BM451184 AGENCOURT
6	1169	53.0	823	7	CO771206 tesc18 ES
7	1162	52.7	779	5	BU447831 BU447831
8	1158	52.5	896	7	CO773091 tesc18 ES
9	1133	51.4	781	7	CN232453 CN232453

10	1107	50.2	932	5	BU541306 AGENCOURT
11	1106	50.1	717	5	BU385761 BU385761
12	1091	49.5	899	2	BE895119 BE895119
13	1079	48.9	772	4	BI562420 BI562420
14	1075	48.7	794	5	BU205357 BU205357
15	1070.5	48.5	834	7	CF593363 CF593363
16	1068.5	48.4	875	4	BI259568 BI259568
17	1063	48.2	888	5	BQ225071 BQ225071
18	1050	47.6	657	4	BG547675 BG547675
19	1050	47.6	682	4	BI333581 BI333581
20	1025	46.5	844	4	BG474616 BG474616
21	1014	46.0	602	7	CO737403 CO737403
22	1005.5	45.6	876	5	BU184346 BU184346
23	990.5	44.9	906	4	BG386905 BG386905
24	968.5	43.9	966	5	BU339278 BU339278
25	967	43.8	631	5	BU112717 BU112717
26	966.5	43.8	1003	5	BQ222020 BQ222020
27	962	43.6	755	5	BU371858 BU371858
28	955	43.3	699	5	BU070784 BU070784
29	944	42.8	604	5	BX261107 BX261107
30	909.5	41.2	716	7	CF745774 CF745774
31	900	40.8	570	5	BU293329 BU293329
32	893	40.5	670	7	CN359673 CN359673
33	889	40.3	928	5	BQ424438 BQ424438
34	880.5	39.9	824	2	BE568528 BE568528
35	870.5	39.5	820	7	CK597557 CK597557
36	862	39.1	672	5	BX261106 BX261106
37	858.5	38.9	1663	6	CD510246 CD510246
38	855	38.8	785	6	CB990192 CB990192
39	850	38.5	523	2	BF2929410 BF2929410
40	828.5	37.6	1063	5	BU356643 BU356643
41	818.5	37.1	743	4	BI562440 BI562440
42	815	36.9	753	7	CK364431 CK364431
43	798	36.2	933	2	BE866040 BE866040
44	786	36.1	986	5	BU751496 BU751496
45	793	35.9	484	5	BX281211 BX281211

ALIGNMENTS

RESULT 1
AK013347
LOCUS
DEFINITION Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length enriched library, clone:281045604 product:kinase interacting with leukemia-associated gene (bractinin), full insert sequence.

ACCESSION AK013347
VERSION AK013347.1 GI:12850651
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Carninci, P., Hayashizaki, Y.
TITLES High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayata, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Normalisation and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)

REFERENCE
AUTHORS Shizata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kasaiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,

|||||
Db 939 AATATTGGCCAGTAAAGCAGTGGGTAATGCGCAATTCAGCCCATCATCCCTCAGAGACCTT 998
|||
Qy 281 TlelySerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCyS 300
|||
Db 999 ATCAAAAGATGATTCATATGATGACCCGGGAGAAAGATCCCTGCTGATGATGGATTTGTGC 1058
|||
Qy 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
|||
Db 1059 AGCCCAATCTTAAAGCATCTCTTTGGCCCTCATATGAAAGATCTGGATGATGCTTCCGACT 1118
|||
Qy 321 ProValLeuArgLeuLeuAsnValLeuAspAspPtyrLeuGlyAsnGluGluGly 340
|||
Db 1119 CCAGTCTCAGGCTCTCTCATATGCTGCTGATGATGATTAACCTTGAATAATGAAATGAAAT 1178
|||
Qy 341 GluAspValValGluAspValValGluGluGlyCysGluGlyPtyrProValValSerLeu 360
|||
Db 1179 GAAGATGTTGTAGAAATGTGAAGAGAGTGTCAAGAAATATGACACAGTGTCTCTCG 1238
|||
Qy 361 LeuValProValGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
|||
Db 1239 CTGTGTCAAAGGAAATCTGACAGAGGCAAGTCTGTGATGATGAGGAAAGCTGGT 1298
|||
Qy 381 AspSerIleAlaAlaGlnIleLeuLeuThrGlyArgMetPheAspGlyLeuPheValVal 400
|||
Db 1299 GATTCCAAAGCTGCTCAGAAAGTTCGACTGGAGAGATGTTGACGGGAAGTTGTGTG 1358
|||
Qy 401 AlaThrPheTyrProLeuSerAlaTyrIleValGlyTyrLeuTyrGlnThrLeuLeu 419
|||
Db 1359 GCTACATTCCTACCCCTGAGTGCCTCAAGAGGGGATATCTTATCAAACTTGTCTT 1415
|||
RESULT 2
AK030152 2846 bp mRNA linear HTC 03-APR-2004
LOCUS
DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched
library, clone:4933401D07 product:kinase interacting with
leukemia-associated gene (stratimin), full insert sequence.
ACCESSION AK030152
VERSION AK030152.1 GI:26326136
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Carninci,P. and Hayashizaki,Y.
1
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Komno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
JOURNAL prepare full-length cDNA libraries for rapid discovery of new genes
MEDLINE 20499374
PUBMED 11042159
REFERENCE
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Komno,H., Akiyama,J., Nishi,K., Kitsuai,T., Taashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Mishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kasahwasgi,K.,
Fujisake,S., Inoue,K., Togawa,Y., Izawa,M., Ozawa,K., Tanaka,T., Wataniki,M.,
Yoneda,Y., Ishikawa,T., Okawa,K., Tanaka,T., Matsura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system-384-format
JOURNAL sequencing pipeline with 384 multicapillary sequencer
MEDLINE Genome Res. 10 (11), 1757-1771 (2000)
PUBMED 20530913
REFERENCE
AUTHORS 4
The RIKEN Genome Exploration Research Group Phase II Team and the

FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
REFERENCE
AUTHORS 5
TITLE The FANTOM Consortium and the RIKEN Genome Exploration Research
JOURNAL Group Phase I & II Team.
MEDLINE Analysis of the mouse transcriptome based on functional annotation
PUBMED of 60,770 full-length cDNAs
6 (bases 1 to 2846)
REFERENCE
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Haneagaki,T., Hara,A., Hashizume,M.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T.,
Hori,F., Imoani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,T., Kojima,Y., Kondo,S., Komno,H., Konda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Nunazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shingawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsunumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@ec.riken.jp,
URL:ftp://genome.gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
COMMENT
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:ftp://genome.gsc.riken.jp/
URL:ftp://fantom.gsc.riken.jp/.
FEATURES
source
1..2846
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:4933401D07"
/db_xref="taxon:10090"
/clone="4933401D07"
/sex="male"
/tissue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
238..1230
/note="unnamed protein product; kinase interacting with
leukemia-associated gene (stratimin) (MGI:1341908)
putative"
/codon_start=1
/protein_id="BAC26812.1"
/db_xref="GI:26326137"
/translation="HTLVGVTTHSPNVPSCRLLELDVSVSELLYSSGQSM
MIQHCARDVLEALFLHNGVHADLKRNTLWSNENCFKIDGSLFKSGQDVKY
IOTDYSRPEAEIOWCLAOAGIOSPTCTSVNDLSGIIILLEMSSGMLKRTVYSOR
WKANSALIDHIFASKAVNAAIPAYHRLDIKSLHDDPGRRIPAEVALCSPPFSIP
PAPHIVEDVIMLPTPLRLNLVDDVYENDEBEVVDVBECKYQPVVSLVPE
NPGROQVDEVYANNADSKAOKLTLGRNFDGKFTVATPTPLSAYRGTLVQTL"
2823..2828
/note="putative"
2846
/note="putative"
ORIGIN
polyA_signal
polyA_site
Alignment Scores:
Pred. No.: 2846
Score: 1711.00 Matches: 326
Percent Similarity: 99.39%

Best Local Similarity: 98.79% Mismatches: 2
Query Match: 77.56% Indels: 0
DB: 3 Gaps: 0

US-10-798-532-4 (1-419) x AK030152 (1-2846)

```
QY 90 ValThrluYrGlyValPheThrIleHisPheSerProAsnValProSerArgCysLeu 109
DB 238 ATGACTTTATATGAGATCTTTACCATACACTTCTCCCAATGTGCATCACCGCTCTCG 297
QY 110 LeuLeuGluLeuLeuAspValSerValSerGluLeuLeuLeuLeuLeuLeuLeuLeu 129
DB 298 CCGCTTGAACCTCTGGATGTCAGTGTTCGAAATTCCTTTATATTCATTCACGAGCC 357
QY 130 CysSerMetTrpMetIleGlnHisCysValArgAspValLeuGluValLeuValPheLeu 149
DB 358 TCGCTCATGTGATGATGATACAGCACTGTGCCGAGATGTTCTGGAGGCCCTCGTTTCTT 417
QY 150 HisHisGluGlyTyrValHisAlaAspLeuLeuProArgAsnIleLeuTrpSerAlaGlu 169
DB 418 CACCATGAGGGCTACCTCATGCAGACCTCAACCAAGAAACATCTGTGGAGTGGGAG 477
QY 170 AangluCysPheLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuVal 189
DB 478 AATGAGTGTTTAAGCTTATGACTTTGACCTCAGCTTCAAGAGGCAATCAGAGCGTA 537
QY 190 LysTyrIleGlnThrAspGlyTyrArgAlaProGluValGluLeuGlnAsnCysLeuAla 209
DB 538 AGTATATTCAGACACACGGGTATCGCGCTCCGAAAGCAGAGCTGCAGAACTGCTGGCC 597
QY 210 GlnAlaGlyLeuGlnSerAspThrGluCysThrSerAlaValAspLeuTrpSerLeuGly 229
DB 598 CAGCGCGGCTGAGAGTGTATCAGATGATCACTCAGCTGTGATCTGTGAGACCTCGGA 657
QY 230 IleIleLeuLeuGlnLeuLeuPheSerGlyMetLeuLeuLeuHisThrValArgSerGlnGlu 249
DB 658 ATCATTTTCTGAAATGTTCTCAGAAATGAACTTAACATACATCAGATTCAGGAA 717
QY 250 TrpLysAlaAsnSerSerAlaIleIleAspHisIlePheHisSerLysAlaValAlaAsn 269
DB 718 TGGAGGCAAAACGTTCTGTCTATTTGATTCATATTTTCCAGTAAGAGTGGTGAAT 777
QY 270 AlaAlaIleProAlaTyrHisLeuAlaGlyAspLeuIleLysSerMetLeuHisAspAspPro 289
DB 778 GCGCAATTCACAGCCATCACTCAGAGACCTTATCAAAAGCATGCTTATGATGACCG 837
QY 290 SerArgArgIleProAlaGluMetAlaLeuCysSerProPhePheSerIleProPheAla 309
DB 838 GGCAGAAAGATCCCTGCTGATGAGATGCGATTTGCAAGCCCATTTCTTACATTCCTTTGCC 897
QY 310 ProHisIleGluAspLeuValMetLeuProThrProValLeuArgLeuLeuValLeu 329
DB 898 CCTCATATTTGAAGATCTGGTGAATGCTTCGACATCCAGTGTCTCCTCATATGTCGTG 957
QY 330 AspAspAspTyrLeuGluAsnGluGluGlyTyrGluAspValValGluAspValLysGlu 349
DB 958 GATGATGATTTACCTTAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 1017
QY 350 GlnCysGluLysTyrGlyProValValSerLeuLeuValProLysGluAsnProGlyArg 369
DB 1018 GAGTGTCAAGAAATATGAGCAAGTGTTCCTCTGCTGCTCCAAAGAAATCTCGCAGA 1077
QY 370 GlnGlnValPheValGluTyrAlaAsnAlaGlyAspSerLysAlaIleGlnLysLeuLeu 389
DB 1078 GGAACAATGCTTCTGATGATCGGAAACGCTGATTCCTCAAGCTGCTCAGAAAGTTCGTG 1137
QY 390 ThrGlyArgMetPheAspGlyLysPheValValAlaThrPheTyrProLeuSerAlaTyr 409
DB 1138 ACTGGAGAGATGTTTGAACGGAAGTGTGTGTGCTTACATTTCAACCGCTGAGTGCCTAC 1197
QY 410 LysArgGlyTyrLeuTyrGlnThrLeuLeu 419
DB 1198 AAGAGGGGATATCTTATCAACCTTGTCT 1227
```

RESULT 3
BM468107 1079 bp mRNA linear EST 05-FEB-2002
LOCUS BM468107
DEFINITION AGENCOURT 6432293 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5535653
5' mRNA sequence.

ACCESSION BM468107
VERSION BM468107.1 GI:16517149
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE NIH-MGC <http://mgs.nci.nih.gov/>.
1 (bases 1 to 1079)
NATIONAL INSTITUTES OF HEALTH, MAMMALIAN GENE COLLECTION (MGC)
UNPUBLISHED (1999)
CONTACT: Robert Strusberg, Ph.D.
Email: cgabbs-remail.nih.gov

Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>
Plate: LLNL12224 row: b column: 06
High quality sequence stop: 754.
Location/Qualifiers
1. 1079

FEATURES
source
1. 1079
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5535653"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb."

ORIGIN

Alignment Scores:
Pred. No.: 3,05e-148 Length: 1079
Score: 1521.50 Matches: 309
Percent Similarity: 92.86% Conservative: 3
Best Local Similarity: 91.96% Mismatches: 17
Query Match: 68.97% Indels: 7
DB: 4 Gaps: 3

US-10-798-532-4 (1-419) x BM468107 (1-1079)

```
QY 90 ValThrluYrGlyValPheThrIleHisPheSerProAsnValProSerArgCysLeu 109
DB 2 GIGACTTTGTATGAGATGTTTACATCCACTTTTCTCCAAATGTGCATCACGCTGTG 61
QY 110 LeuLeuGluLeuLeuAspValSerValSerGluLeuLeuLeuLeuLeuLeuLeuLeu 129
DB 62 TTGCTTGAACCTCTGGATGTCAGTGTTCGAAATGCTTATATTCAGTCACACAGGT 121
QY 130 CysSerMetTrpMetIleGlnHisCysValArgAspValLeuGluValLeuValPheLeu 149
DB 122 TGTTCATGTGATGATATACAGCATTTGTGCCCAGAGATGTTTGGAGGCCCTTCTTTCTT 181
QY 150 HisHisGluGlyTyrValHisAlaAspLeuLeuProArgAsnIleLeuTrpSerAlaGlu 169
DB 182 CATCATGAGGCTATGTTCATCGCACTCAACACGTAACATATTTGTGGATGTCAGAG 241
QY 170 AangluCysPheLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuVal 189
DB 242 AATGATGTTTAAACTCATTTGGAATTTGGAATTTGGAATTTGGAATTTGGAATTTGGA 301
QY 190 LysTyrIleGlnThrAspGlyTyrArgAlaProGluValGluLeuGlnAsnCysLeuAla 209
```

```

Db 302 AAGTATATTCAGACACACGGGTATCCGGCTCCAGAAAGCAGAAATTCGAAATTCCTTGCC 361
Qy 210 GlnAlaGlyLeuGlnSerAspThrGluCysThrSerAlaValAspLeuTPSPSerLeuGly 229
Db 362 CAGGCTGGCCCTGAGAGTGTATACAGATGACTCAGCTGTATATCTGTGGACCTTAGGA 421
Qy 230 IleIleLeuLeuGlnMetPheSerGlyMetLeuLeuYehiSphrValArgSerGlnGlu 249
Db 422 ATCATTTTATCGAATATGTTTCTCAGCAATGAACCTGAACATCAGTCAGATTCAGGAA 481
Qy 250 TTPYAlaIaenSerSerAlaIleIleAspHisIlePheAlaSerIlyAlaValIaen 269
Db 482 TGGAAAGCAAAACAGTTCTGCTATTAATGATCATATTTTCCAGTAAGACAGTGGAAT 541
Qy 270 AlaAlaIleProAlaThrIleLeuArgAspLeuIleLeuSerMetLeuHisAspAppro 289
Db 542 GCGCAATTCACAGCTATCAGCTTAAGAGACCTTATCAAAAGCATGCTTCATGTATGATCCA 601
Qy 290 SerArgArgIleProAlaGluMetAlaLeuCysSerProPheSerIleProPheAla 309
Db 602 AGCAGAAAGATTCCTCTCTGAAATGGCATTTGTGAGCCCATCTTTAGCATCTTTTGGC 661
Qy 310 ProHisIleGluAspLeuValMetLeuProThrProValIleuArgLeuAsnValLeu 329
Db 662 CCTCATATTGAAGATCTGCTCATGCTTCCACATCCAGTCTAAGACTGCTGAATGTGCTG 721
Qy 330 AspAspAspArgIleuGluAsnGluGluGluGluGluGluGluGluGluGluGluGlu 349
Db 722 GATGATGATGATTTATTTGAGATGAGAGGATGAGATGAGATGAGATGAGATGAGATGAG 781
Qy 350 GluCysGlnIlyrTyrgIyProValIaSerLeuLeuValProlyGluAsnProGlyArg 369
Db 782 GAGTGTCAAAAATATGACACAGTGGTATCTCTCTCTGATTCAGAAAGAAATCTTGCCAG 841
Qy 370 GlyGlnVal-PheValGluTyraIaenAla-GlyAspSerIlyAlaIaGlnIlyLeu 388
Db 842 GGCAGAGTCTTTGNTGAGATGACAAATGCTGNGTATTCAAAAGCTGGCAGAAATTTA 901
Qy 389 LeuThrGlyArg---MetPheAspGlyIyPheValIaIa-ThrPheTyrrProLeu-- 406
Db 902 CTGACGTGGAANAAGATGGTTGATGGGAAAGTTGNTGTTGTCCTCAACAATCTCACTG 961
Qy 407 -SerAlaTyrrIyAspGly---TyrrLeuTyrgInThrLeu 418
Db 962 GAGTGTCTTACAGAGAGGAGTATTCTGTATTCAAACCTT 1001

RESULT 4
BY12344 1015 bp mRNA linear EST 17-DEC-2002
LOCUS BY12344
DEFINITION Mus musculus cDNA clone 2810454004 5', mRNA sequence.
ACCESSION BY12344
VERSION BY12344.1 GI:27123621
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus.
1 (bases 1 to 1015)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Oasato, N., Saito, R., Suzuki, H., Yamana, I.,
Kiyosaki, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojodori, T., Balderelli, R., Hill, D.P., Bult, C.,
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K.W., Blake, J.A., Brad, D., Bruscia, V.,
Choitha, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
Fletcher, C.P., Forrest, A., Frazer, K.S., Gasterland, T.,
Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
Gustinchin, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.U., Konagaya, A.,
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,

```

Numata, K., Okido, T., Pavan, M.J., Pertea, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramchandran, S.,
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
Sandelin, A., Schneider, C., Semple, C.A., Secou, M., Shmida, K.,
Sultana, R., Takemata, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, T.,
Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavalan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imocani, K., Ishii, Y.,
Itoh, M., Kagawa, T., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsunuma-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P.,
Fukuda, S., Hashizume, W., Hayashizaki, K., Hirozane, T., Hori, F.,
Imocani, K., Ishii, Y., Itoh, M., Kagawa, T., Kawai, J., Kojima, Y.,
Kondo, S., Kono, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M.,
Nomura, K., Nunazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N.,
Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,
Takeda, Y., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission

Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multipipillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.

FEATURES

source

Location/Qualifiers

1..1015
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/cclone="2810454004"
/tissue_type="whole body"
/dev_stage="10, 11 days embryo"
/clone_lib="RIKEN full-length enriched, 10, 11 days embryo
whole body"

ORIGIN

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:
Score: 1271.50	1015	251	5
Percent Similarity: 89.51%			
Best Local Similarity: 87.76%			
Query Match: 57.64%			
DB: 6	Gaps:	4	

US-10-798-532-4 (1-419) x BY712344 (1-1015)

QY 1 MetAlaGlySerGlyCysValATrpgAlaGluProProkrpPheLeuGluAlaPheGly 20

Db 159 ATGGCGGGGCTCCGGCTGGCGGGGCGCGAGCCCGCGCTTCTTGAGGCGCTTCGGG 218

QY 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40

Db 219 CGGCTGTGGAGGCTCCAGAGCCGCTGGGAGCGGCTCGTGGCTCGGCTGACCGGGTG 278

QY 41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60

Db 279 CGGTGCTGGGTACCCCGGGCTGGCCCGCGGCGCTCAAGCAGTTCCTGCTCCGGGA 338

QY 61 ThrThGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgGlyGluArgAlaAlaLeu 80

Db 339 ACCACCGGGGCTCCGGCTTCGGCGGAGTATGTTTCCGCAAGAGAGGCGGCGCTG 398

QY 81 GlnGlnLeuGlnGlnValArgAsnIleValThrLeuTyrGlyValPheThrIleAlaPhe 100

Db 399 GANCACTTGACAGGCTCAGAGAACATCGTGACTTTATATGAGTCTTTACATACACTTC 458

QY 101 SerProAsnValProSerArgCysLeuLeuLeuGlnLeuLeuAspValSerValSerGln 120

Db 459 TCTCCCAATGTGCATCAGCTGTCTGTCTGCTTGAACCTCTGATGTCACTGTTTGGAA 518

QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140

Db 519 TTGCTTTATATTCCAGTCACTCAGGCGTCTCATGTGATGATGATACAGCATGTGCCCA 578

QY 141 AspValLeuGlnAlaLeuAlaPheLeuHisGlnGlyTyrValHisAlaAspLeuLys 160

Db 579 GATGTTCTGGAGGCCCTCGCTTTTCTTACATGAGGCGTACGTCATGACAGCTCAAA 638

QY 161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180

Db 639 CCACGAAACATCTGTGAGTGTGGAGAAAGATGTTTAAAGTTTATGACTTGGACTC 698

QY 181 SerPheLeuGlnGlyValAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200

Db 699 AGCTTCAAGAGAGGCAATCAGACGTAAGTATATTCAAGCAGACGGGATTCGGCTCCT 758

QY 201 GlnAlaGlnLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGlyCysThr 220

Db 759 GAAGCAGAGCTGCAGAACTGCTTGCCCGAGGCGGCTGCAGATGATACAGATGTACC 818

QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnMetPheSerGlyMetLys 240

Db 819 TCGCTGTGTATGTCTCGGGCTCGGAATCCTTTACTGTGAGATGTTCTCAGGAATGAAG 878

QY 241 LeuLys---HisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAsp 259

Db 879 CTGAATAATACATCAGATCTCAGAG-----TGGGAGGCA---AGCAGTCTGCATCAATTGC 929

QY 260 HisIlePheAlaSerLysAlaValAlaAsnAlaIleIleProAlaTyrHisLeuArgAsp 279

Db 930 CATATATTT-----GCAGTGACAGAGTGTGTTCGCAATTCAGCCATCACTCAGAGC 980

QY 280 LeuIleLysSerMetLeu 285

Db 981 GCTTATCAGAGCATGCTC 998

RESULT 5

LOCUS BM451184 917 bp mRNA linear EST 05-FEB-2002

DEFINITION AGENCOURT_6392719 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5495159

ACCESSION BM451184

VERSION BM451184.1 GI:18500224

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE

AUTHORS NIH-MGC http://mgi.nci.nih.gov/

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cga@bs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Agencourt Bioscience Corporation

Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

plate: L1M12121 row: 3 column: 24

High quality sequence stop: 718.

Location/Qualifiers

1..917

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5495159"

/tissue_type="retinoblastoma"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_67"

/note="Organ: eye; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo Average insert size 1.75 kb. Library constructed by Life Technologies."

ORIGIN

Alignment Scores:

Pred. No.: 4,47e-119 Length: 917

Score: 1242.00 Matches: 240

Percent Similarity: 99.59% Conservative: 0

Best Local Similarity: 99.59% Mismatches: 1

Query Match: 56.30% Gaps: 0

DB: 4 Indels: 0

US-10-798-532-4 (1-419) x BM451184 (1-917)

QY 179 GlnLeuSerPheLysGlnGlyValAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArg 198

Db 22 GGGCTTAAGCTTCAAGAGGCAATCAGATGTAAATATTTCAGACAGCGGTATCGG 81

QY 199 AlaProGlnAlaGlnLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGlu 218

Db 82 GCTCCAGAGCAGAAATGCGAAATGCTTGCGCCAGGCTGGCTCAGAGTATACAGAA 141

QY 219 CysThrSerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnMetPheSerGly 238

Db 142 TGAACCTCAGCTGTGATGTGTGAGGCTGAGAAATCATTTTACTGGAATGTTCTCAGGA 201

QY 239 MetLysLeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIle 258

Db 202 ATGAAGCTGAACATACATCAGATCTCAGAGATGAAGGCAAGAGTTCTGCTATATT 261

QY 259 AspHisIlePheAlaSerLysAlaValAlaAsnAlaIleIleProAlaTyrHisLeuArg 278

Db 262 GATCACATATTGCGCAGTAAAGCATGTGATGCCGCAATTCAGCCATACACTAAGGA 321

QY 279 AspLeuIleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGlnMetAla 298

Db 322 GACCTTATCAAAAGCATGCTCATGATGATCAAGCAGAGAAATTCGTGTAATGCGA 381

QY 299 LeuCysSerProPhePheSerIleProPheAlaProHisIleGlnAspLeuValMetLeu 318

Db 382 TTGTGAGGCCCATTTCTTACATTCCTTTTGGCCCTCATATTGAAGATCTGCTCATGCTT 441

QY 319 ProThrProValLeuArgLeuLeuAsnValIleAspAspAspTyrLeuGlyValAsnGln 338

Db 442 CCACCTCCAGTCTAAGACTGCTGAATGTCTGAGATGATGATTAATCTTGAGAAATGAAG 501

QY 339 GtUTYrGLuAspValValGluAspValLysGluGluCysGluLysTyrGlyProValVal 358
| | | | |
Db 502 GAAATATGAAGATGTTTACAGATGTAAAGAGGAGTGTCAAAAATATGAGCACAGTGTAA 561
| | | | |
QY 359 SerLeuLeuValProLysGluAspProGlyValGlyGlnValPheValGluTyrValAaen 378
| | | | |
Db 562 TCTCACTGTTCCAAAGGAAATCTCGCAGAGGCAAGTCTTTTCAATATGATGCAAAAT 621
| | | | |
QY 379 AAGLAspSerLysValAaLagLysLeuLeuThGlyAArgMerPheAspGlyLysPhe 398
| | | | |
Db 622 GCGGAGATTCAGAACTGGCAGAAATTACTGACTGAAAGATGTTTATGGGAAATTT 681
| | | | |
QY 399 ValValAaLethPheTyrProLeuSerAlaTyrLysAArgGlyTyrLeuTyrGlnThrLeu 418
| | | | |
Db 682 GTTGTGCGTCAACATTAACCGCTGAGTGTGCTTCAAGAGGGGAAATATCTGATCAAACTTG 741
| | | | |
QY 419 Leu 419
| | |
Db 742 CTT 744
| | |
RESULT 6
LOCUS CO771206 823 bp mRNA linear EST 04-AUG-2004
DEFINITION testis cDNA library Gallus gallus CDNA 3', mRNA
sequence.
ACCESSION CO771206
VERSION CO771206
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 823).
Shin,J.
TITLE Testis cDNA Library
JOURNAL Unpublished (2004)
COMMENT Contact: JiHye Shin
Division of Animal Genetic Engineering
School of Agricultural Biotechnology, Seoul National University
San 56-1, Sillim-dong, Gwanak-gu, Seoul, 151-742, Korea
Tel: +82-31-479-7557
Fax: +82-31-479-7550
Email: paecljoh@hanmail.net
Seq primer: T3
FEATURES
source Location/Qualifiers
1..823
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Korean Native chicken"
/db_xref="taxon:9031"
/sex="Male"
/tissue_type="Testis"
/dev_stage="26weeks"
/clone_lib="Testis cDNA library"
/note="Vector: Uni-ZAP XR vector; Site_1: XhoI; Site_2:
EcoRI; The libraries were synthesized using the
ZAP(R)-cDNA synthesis method (Stratagene). cDNA was
prepared, size-fractionated and inserted into the Uni-ZAP
XR vector using XhoI linker-primer and EcoRI adaptor.
After in vivo excision with B. coli strain SOLR, cDNA
libraries from testis contain each insert size range of
0.5-3kb (n=20), and 0.5-3.5kb (n=20)."

ORIGIN
Alignment Scores:
Pred. No.: 1,68e-111 Length: 823
Score: 1169.00 Matches: 229
Percent Similarity: 84.38% Conservative: 14
Best Local Similarity: 79.51% Mismatches: 23
Query Match: 52.99% Indels: 22
DB: 7 Gaps: 2

US-10-798-532-4 (1-419) x CO771206 (1-823)
QY 47 GlySerProPro-----GlyValLeuArgGlnPheLeuProProGlyThrThrGly 63
| | | | |
Db 6 GAGAGCTCAGACCGCGGTGGCCNGGCCCTTACAACTAGTGAATCCCCCGGCTGCAGAAAT 65
| | | | |
QY 64 ALaAaLaAseRAlaAaLagLysLysPheArgLysGluArgAlaAaLeuGluLeu 83
| | | | |
Db 66 TCGGCACGAGGC----- 77
| | | | |
QY 84 GlnGlyHlaArgAenLLeValThrLeuTyrGlyValPheThrLLeHlaPheSerProAsn 103
| | | | |
Db 78 -----CGCAACATGTGATCTCTGACCGCGGTTCACCAACCACTACTCGGCCAAC 128
| | | | |
QY 104 ValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGluLeuLeu 123
| | | | |
Db 129 GGCCTGCTCCGCTGCTGCTGCTGAGCTGTGATACAGGATCTGAGCTGCTGCTG 188
| | | | |
QY 124 TyrSerSerHlaGlnGlyCysSerMetTrpMetLLeGlnHlaCysAlaArgAspValLeu 143
| | | | |
Db 189 CACTCCAGCAACCAAGGCGTGTCCATGTGATGATCCAGCACTGCGCGGAGCGTGTG 248
| | | | |
QY 144 GlnAlaLeuAlaPheLeuHlaHlaGluGlyTyrValHlaAaPLeuLysProArgAsn 163
| | | | |
Db 249 GAAGCCCTGGCCTTCTCTGCACCAAGGCTACGACGACGACCTGAAGCCACGCAAC 308
| | | | |
QY 164 LLeLeuTTPSerAlaGluAsnGluCysPheLysLeuLLeAspPheGlyLeuSerPheLys 183
| | | | |
Db 309 ATCTGTGTGAGGCGCAGAGAGAGTGTAAAGCTCATTTGACTTGAAGCTTAACTCAAA 368
| | | | |
QY 184 GlnGlyAsnGlnAspValLysTyrLLeGlnThrAspGlyTyrArgAlaProGluAaLagLu 203
| | | | |
Db 369 GAGGGGAATCAGATGTGAATATTTTCAACAGCGGATTCGGGCTCCAGAGCAGAA 428
| | | | |
QY 204 LeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThrSerAlaVal 223
| | | | |
Db 429 CTGCAGAACTGCTGACACAGGCAAGGCTCCAGATGAGACGAAATGATCACTGCTGTG 488
| | | | |
QY 224 AspLeuTTPSerLeuGlyLLeLLeuLeuGlnMetPheSerGlyMetLysLeuLysHla 243
| | | | |
Db 489 GATCTGTGTGAGTGTGGAATGTTTACTGAAAGTGTCTCAGGAATATACTGAACAT 548
| | | | |
QY 244 ThrValArgSerGlnGlyTTPlysAlaAsnSerSerAlaLLeLLeAspHlaLLePheAla 263
| | | | |
Db 549 ACGATCCAAATCTCAGGAATGGAAGCAAACTTGTGCATCATGATCGCATTTTTCG 608
| | | | |
QY 264 SerLysAlaValValAsnAlaLLeLLeProAlaTyrHlaSLeuArgAspLeuLLeLysSer 283
| | | | |
Db 609 AGTGAAGGGGTGTAAATTAATCAAGCCATTCAGCTTATCACTCAGAGACTTATTAAGC 668
| | | | |
QY 284 MetLeuHlaAspAspProSerArgArgLLeProAlaGluMetAlaLeuCysSerProPhe 303
| | | | |
Db 669 ATGCTTCATTTGTGACCAAGCAGACCTCTGTGAAAGGCTTTATGAGCCCATTC 728
| | | | |
QY 304 PheSerLLeProPheAlaProHlaLLeGluAspLeuValMetLeuProThrProValLeu 323
| | | | |
Db 729 TTCAGCATTCCTTTGCTCCCATATTTGAAGATTTGGATGCTCCACAGCGCTGTGCTG 788
| | | | |
QY 324 ArgLeuLeuAsnValLeuAspAsp 331
| | | | |
Db 789 AGCTGCTGAATGTTCTTAACGAT 812
| | | | |
RESULT 7
LOCUS BU447831 779 bp mRNA linear EST 29-NOV-2002
DEFINITION 603767571F1 CSQRBN13 Gallus gallus CDNA clone ChEST695n17 5', mRNA
sequence.
ACCESSION BU447831 GI:25937142
VERSION BU447831
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianidae; Gallus.
 1 (bases 1 to 779)
 Boardman, P. B., Sanz-Eguero, J., Overton, I. M., Burt, D. W., Boech, E.,
 Fong, W. T., Tickle, C., Brown, W. R. A., Wilson, S. A. and Hubbard, S. J.
 A Comprehensive Collection of Chicken cDNAs
 Curr. Biol. 12 (22), 1965-1969 (2002)
 2235534
 MEDLINE
 PUBMED
 12445392
 Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology
 (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.
 Location/Qualifiers
 1..779
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="Layer"
 /db_xref="taxon:9031"
 /clone="CHEST695n17"
 /sex="Female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_1lb="CSEORBN13"
 /note="Organ: ovary; Vector: pBluescript II KS(+); Site_1:
 EcoRI; Site_2: NotI; This normalized library was
 constructed from 1 million independent clones. cDNA
 synthesis was initiated using an oligo(dT) primer, using
 methylated C in the first strand synthesis reaction.
 Following this first strand reaction, double-stranded cDNA
 was blunt-ended, ligated to NotI adapters, digested with
 EcoRI, size-selected, and cloned into the NotI and EcoRI
 compatible sites of a custom modified MCS of the
 pBluescript (KS+) vector. The library was normalized in 2
 rounds using conditions adapted from Soares et al., PNAS
 (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
 (1996): 791, except that a significantly longer
 reannealing hybridization was used."

ORIGIN
 Alignment Scores:
 Pred. No.: 8,36e-111 Length: 779
 Score: 1162.00 Matches: 224
 Percent Similarity: 93.02% Conservative: 16
 Best Local Similarity: 86.82% Mismatches: 18
 Query Match: 52.67% Indels: 1
 DB: 5 Gaps: 0

US-10-798-532-4 (1-419) x BU447831 (1-779)
 QY 158 Asplleuylsefrolrganilleleutpseralgluabnglucyphelyseuileap 177
 Db 6 GACCTGAAGCAGCAACATCTCTGTGAGCGCGAGAGAGAGTCTTAAGCTCATTTGAC 65
 QY 178 pnegllyleuserthelysgluygaenglnasvallyetyrtlllegrnthrapgllytyr 197
 Db 66 TTTTGACTTAAGCTTCAAGAGGGAATCAGAGATGTGAAATATTTTCAACACAGCGGTAT 125
 QY 198 ArgAlProgluagluuenglnaenCyseuAlgluAlaglyleuenglnaseAPTtr 217
 Db 126 CGGGCTCCAGAGGCAAGACTGCGAAGCTGCTTGCACAGCAGCGCTCCAGATGAGAGC 185
 QY 218 glucyethseralavalaspleutpserleuglylleleuengluwetpheser 237
 Db 186 GAATGACCTCTCTGTGATCTGTGAGTCTGGAATTTTAACTCGAATATTTCTCA 245
 QY 238 GYMeLyleuLysHsthrValArgSerGlnGluTrrplyAlaenSerSerAlaIle 257

Db 246 GGAATGAAATGAAACATACCTCAATCTCAGGAATGAGACAAACAGTTCCGCATC 305
 QY 258 lileaphisilephealaserlysalavalalaenalalaleproalatyrtlsreu 277
 Db 306 ATCGATCGATTTTGGCCAGTGAAGGGGTGTTATTCAGCATTCAGCTTATACCTC 365
 QY 278 ArgapleuilelysermetleuHlsapappproserrhaglylleproAlaIleuMet 297
 Db 366 AGAGACTTATTAAGACATGCTTCATTTGTGACCAAGCAGCGAGCTGCTGTAAG 425
 QY 298 AlaleuCyesserPropheserlleProphelaprophisillegluAspleuValMet 317
 Db 426 GCTTATGAGCCATCTTCTTCAGCATTCCTTGTGCTCCCATATGAAGATTCGTTAG 485
 QY 318 lcuProthProvalleuArgleuLeuAnValleuAspAspTyrtleuglyyansglu 337
 Db 486 CTCCCAAGCCCTGTCTGAGCTGCTGAATGTTCTTAAGCATGCTTCTCTGACAGTGA 545
 QY 338 gluglutyrgluaspvalaigluaspvallysgluyglucyeglnlyetyrglyProval 357
 Db 546 GAAGAAATAGAAATATCTCTGAGAACATTAAGCGAGGTGTCAAGAAATATGACCGGTG 605
 QY 358 ValserleuLeuValProlysgluAnProgluyArglyglnValPheValglutyryla 377
 Db 606 GTTCTCTGCTTATTCGAGAGAAATCTGTAAAGGCCAAAGCTTTGT-GAATATGCA 664
 QY 378 AenAlaglyAspSerlyeAlaAlaglnlyseuLeuThrgluyArgMetPheAspelylys 397
 Db 665 AATGCTGTGATTCGAAAGCTGCCAAAGAAATGCTGACGTGGAAGATTTTGTATGCAAG 724
 QY 398 PheValAlaAlaThrPheTytrProleuserAlatyrylbaaglytyrleuty 415
 Db 725 TCTGTCGTGCTACGNTTACCACCTGAGTGTCTTAAGAGAGATATCTGTAC 778

RESULT 8
 LOCUS CO773091 896 bp mRNA EST 04-AUG-2004
 DEFINITION testis cDNA library Gallus gallus cDNA 3', mRNA
 sequence.
 ACCESSION CO773091
 VERSION CO773091.1 GI:50974358
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianidae; Gallus.
 1 (bases 1 to 896)
 Shin, J.
 Testis cDNA library
 Unpublished (2004)
 Contact: JiHy Shin
 Division of Animal Genetic Engineering
 School of Agricultural Biotechnology, Seoul National University
 San 56-1, Shillim-dong, Gwanak-gu, Seoul, 151-742, Korea
 Tel: +82-31-479-7557
 Fax: +82-31-479-7550
 Email: paecljo@hanmail.net
 Seq primer: 73.
 Location/Qualifiers
 1..896
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="Korean Native chicken"
 /db_xref="taxon:9031"
 /sex="Male"
 /tissue_type="Testis"
 /dev_stage="26weeks"
 /clone_1lb="Testis cDNA library"
 /note="Vector: Uni-ZAP XR vector; Site 1: XhoI; Site 2:
 EcoRI; The libraries were synthesized using the
 ZAP(R)-cDNA synthesis method (Stratagene) cDNA was
 prepared, size-fractionated and inserted into the Uni-ZAP

XR vector using XhoI linker-primer and EcoRI adaptor.
After in vivo excision with E. coli strain SOLR, cDNA
libraries from testis contain each insert size range of
0.5-3kb (n=20), and 0.5-3.5kb (n=20)."

ORIGIN

Alignment Scores:

Pred. No.:	2,71e-110	Length:	896
Score:	1158.00	Matches:	240
Percent Similarity:	81.59%	Conservative:	17
Best Local Similarity:	76.19%	Mismatches:	31
Query Match:	52.49%	Indels:	27
DB:	7	Gaps:	2

US-10-798-532-4 (1-419) x CO773091 (1-896)

```

OY 46 ProGlySer-Propio-----GlyAlaLeuArgGlnPheLeuProGlyThrTh 62
DB 11 CCGGNACCTCCACCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 70
OY 62 rglYAlaAlaSerAlaAlaGluTyrGlyPheArgGlyGluArgAlaAlaLeuGlu 82
DB 71 GAATTCGGACGAGGC----- 86
OY 82 nleuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPheSerPr 102
DB 87 -----CGCACATCTGTACTCTGTAAGCGGCTTACACACACACTCTGGC 133
OY 102 oAnValProSerArgCyLeuLeuLeuGluLeuLeuAspValSerValSerGluLeu 122
DB 134 CAACGCGCGGCTCCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 193
OY 122 uleuTyrSerSerHisGlnGlyCySerSerMetTrpMetIleGlnHisCyAlaArgAspVa 142
DB 194 GCGCACTCCACGACCAAGGCGCTGCTCACTGTAAGATTCACACACTGCGCGCGACG 253
OY 142 lleuGlnAlaLeuAlaPheLeuHisIleGluGlyTyrValHisAlaAspLeuYsProAr 162
DB 254 GCGGAGAGCCCTGGCTTCTCTGCAACAAGGCTACGTCCAGCAGACTGAGACG 313
OY 162 gAnIleLeuTyrSerAlaGluAsnGluCyPheLeuLeuAspPheGlyLeuSerPh 182
DB 314 CAACATCTGCTGAGCGCCGACGAGAGGAGCTTAAAGCTCATGACTTGGACTTACCTT 373
OY 182 elySgLuGlyAsnGlyAspValYsTyrIleGlnTrpAspGlyTyrArgAlaProGlnAl 202
DB 374 CAAGAGGGGAAATCAGATGTGAATATATTAACAACAGAGGAGTATCGGCTCCAGAGGC 433
OY 202 agluLeuGlnAsnCyLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCyThrSerAl 222
DB 434 AGNACTGCAATCTGCTTACACAGGCGGCTCCAGAGTGAACGGAATGTACCTCTGC 493
OY 222 aValAspLeuTyrSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetYsLeuY 242
DB 494 TGTGATCTGTGAGTCTGGGAAATGTTTAACTGAAATGTTCTCAGAAATGAATGAA 553
OY 242 sHisThrValArgSerGlnGluTyrYsAlaAsnSerSerAlaIleIleAspHisIlePh 262
DB 554 ACATACAGTCCAAATCTCAGAAATGAAGACAAACAGTTCGCATATCGATGCAATTTT 613
OY 262 eAlaSer-LysAlaValAlaAsnAlaAlaIleProAlaTyrHisIleuArgAspLeuIle 282
DB 614 TGCATGTAAGGGGGGTGTGTTAATTCAGCCATTTCAGCTTATCCTCAGAGACTTATTA 673
OY 282 ySerSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCySerP 302
DB 674 AAAGCATGTCTTATGTACCAAGCAAGCAGCCTCTGCTGAAAGGCTTATGACGCC 733
OY 302 roPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThrProY 322
DB 734 CATTTTCAGCATTCCTTTGCTCCCAATATGAAAGTTTGTGATGCTCCCAAGCCTGT 793
OY 322 alLeuArgLeu-LeuAnValLeuAspAspTyrLeuGlyAsn-GluGluGluTyrG 341

```

```

DB 794 TCCTAGAGCTGACGTAATCTTCAAGCATGCTTCTGACAGAGTGAAGAAATACGA 853
OY 341 uAspValValGlu-AspValYsGluGluCyGln 352
DB 854 AGAATCCCTGAGACATAGGAGAGAGATGTTCAG 888

```

RESULT 9

CN32453

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 781)
Savolainen, P., Fitzsimmons, C.J., Arvestad, L., Andersson, L. and Lundberg, U.
EST analysis of brain and testis cDNA libraries from White Leghorn and Red Jungle Fowl
Unpublished (2004)
Contact: Peter Savolainen
Department of Biotechnology
Royal Institute of Technology, KTH
SE-106 91 Stockholm, SWEDEN
Tel: +46 (0) 8 5537 8481
Fax: +46 (0) 8 5537 8335
Email: Peter.Savolainen@biotech.kth.se
Seq primer: M13 reverse primer.
Location/Qualifiers
1..781
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn"
/db_xref="taxon:9031"
/sex="male"
/lab_host="ElectroMAX DH10B (Invitrogen)"
/clone_host="Wttestis"
/note="Organ: testis; Vector: pSPORT-1; Site 1: Hind III; Site 2: EcoRI; The cDNA libraries were created with the Superscript Plasmid System (Invitrogen)."

FEATURES

source

ORIGIN

Alignment Scores:

Pred. No.:	9.14e-108	Length:	781
Score:	1133.00 <td>Matches:</td> <td>225</td>	Matches:	225
Percent Similarity:	93.10% <td>Conservative:</td> <td>18</td>	Conservative:	18
Best Local Similarity:	86.21% <td>Mismatches:</td> <td>3</td>	Mismatches:	3
Query Match:	51.36% <td>Indels:</td> <td>0</td>	Indels:	0
DB:	7	Gaps:	0

US-10-798-532-4 (1-419) x CN32453 (1-781)

```

OY 109 leuLeuLeuGlnLeuLeuAspValSerValSerGluLeuLeuTyrSerSerHisGln 128
DB 2 CCGCTGCTGAGACTGCTGATATTCAGCGGTCTGAGCTGCTGCTCAGCTCCGCAACG 61
OY 129 GlyCySerMetTrpMetIleGlnHisCyAlaArgAspValLeuGluAlaPhe 148
DB 62 GGTGTCTCATGTGATGATTCAGCACTGCGCGCGACGCTCTGAAAGCTTGGCTTC 121
OY 149 leuHisIleGluGlyTyrValHisAlaAspLeuYsProArgAsnIleLeuTyrSerAla 168
DB 122 CTGCACCAACAAGGCTAGCTGACGACGCTGAAAGCCACGCAATCTGTGTGAGCGCG 181
OY 169 GluAsnGluCyPheLeuYsLeuIleAspPheGlyLeuSerPheYsGluGlyAsnGlnAsp 188
DB 182 GAGAGAGTGTCTTAAAGCTCATGACTTGTGACTTACCTTCAAGAGGGGAAATCAGAT 241

```


DB 869 CAGGAGATG-----GAAACTGAGAAACTACCGGTCCAGA 904
QY 257 |||::: 11e1le 258
DB 905 ATCTTC 910

RESULT 11
BU385761 717 bp mRNA linear EST 28-NOV-2002
LOCUS BU385761
DEFINITION 60358260F1 CSEQCHN75 Gallus gallus cDNA clone CHEST533d13 5', mRNA
sequence.
ACCESSION BU385761
VERSION BU385761.1 GI:25893762
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
1 (bases 1 to 717).
Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.,
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
22335534
PUBMED 12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers

FEATURES
source
1..717
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White leghorn, H1sex"
/db_xref="taxon:9031"
/clone="CHEST533d13"
/dev_stage="36"
/lab_host="DH10B"
/clone_1lb="CSEQCHN75"
/note="Organ: trunks; Vector: pBluescript II KS(+);
Site_1: EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunt-ended, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reamealing hybridization was used."

ORIGIN
Alignment Scores:
Pred. No.: 5.41e-105 Length: 717
Score: 1106.00 Matches: 210
Percent Similarity: 93.72% Conservative: 14
Best Local Similarity: 87.87% Mismatches: 14
Query Match: 50.14% Indels: 1
DB: 5 Gaps: 0

US-10-798-532-4 (1-419) x BU385761 (1-717)

QY 134 MetTtGlnHtCysAlaArgAspValLeuGluAlaLeuAlaPheLeuHtShiGlnGly 153
DB 3 ATGATCCAGACACGCGCCGCGACGTCCTGGAAGCCCTTGCCCTGCACCAAGGC 62

QY 154 TyrValHtSaLaAspLeuLeuProArgAsnIleLeuTTPSerAlaGluAsnGluCysPhe 173
DB 63 TACGTGACGACGACCTGAAGCCAGCAACATCCTGTGAGGCGCGAGAGAGCTTT 122

QY 174 LysLeuIleAspPheGlyLeuSerPheLysGlnGlyAaGlnAspValLysTyrIleGln 193
DB 123 AAGCTCATTTGACTTGGACTTGTAGCTTCAAGAGGGGAATTCAGAGATGTAAATATTTCAA 182

QY 194 ThrAspGlyTyrArgAlaProGluAlaGluLeuGlnAsnGlyLeuAlaGlnIleLeu 213
DB 183 ACAGACGGGTATCGGCTCCAGAGGCAAGACTGCAAACTGCTAGCAAGGCAAGGCTC 242

QY 214 GlnSerAspThrGlnCysThrSerAlaValAspLeuTTPSerLeuGlyIleLeuLeu 233
DB 243 CAGAGTGACAGCGAATGTATACCTCTGCTGTGATCTGTGAGCTGTGGAATGTTTACTG 302

QY 234 GluMetPheSerGlyMetLysLeuLysHtThrValArgSerGlnLutTyrAlaAsn 253
DB 303 GAAATGTTCTCAGAAATGAACCTGAACATACATCAATCTCAGAAATGGAACAAAC 362

QY 254 SerSerAlaIleIleAspHisIlePheAlaSerLysAlaValAlaAsnAlaIlePro 273
DB 363 AATTCTGCCATCATGATCGCATTTTTCACAGTGAAGGCGGTGTTAATTCAGCCATTCCA 422

QY 274 AlaTyrHisLeuArgAspLeuIleLysSerMetLeuHisAspAspProSerArgTyle 293
DB 423 GCTTATACCTCAGACACTTATTAAGCATGCTTCTTGTGACCAAGCAAGCAAGGC 482

QY 294 ProAlaGluMetAlaLeuCysSerProPhePheSerIleProPheAlaProHisIleGlu 313
DB 483 TCTGCTGAAAAGCTTTATGACGCCATTCCTTCAACATTCCTTGTGCTCCCATTTGAA 542

QY 314 AspLeuValMetLeuProThrProValLeuArgLeuLeuAsnValLeuAspAspTyr 333
DB 543 GATTGGAGATGCTCCCAAGCTGTGCTGAGCTGCGAATGTTCTAAGCATCTTCT 602

QY 334 LeuGlyAsnGluGluGlyTyrGluAspValValGluAspValLysGluCysGlnLys 353
DB 603 CTGCAGAGTGAAGAAAGATACGAATATTCCTGGAAGCATATAGGAGAGAGTGTGAAA 662

QY 354 TyrGlyProValValSerLeuLeuValProLysGluAsnProGlyArgGlyGln 371
DB 663 TATGAGACCGGTGTTCTTCTTATTCGAAAGAAACATCTGTGTAAGGCCAA 717

RESULT 12
BE895119 899 bp mRNA linear EST 20-OCT-2000
LOCUS BE895119
DEFINITION 601436041F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3920939 5',
mRNA sequence.
ACCESSION BE895119 GI:10358193
VERSION BE895119.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 899)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/BLAN at:
<http://image.llnl.gov>
Plate: LMA9753 row: b column: 12
High quality sequence stop: 626.
Location/Qualifiers

FEATURES

source

1. .899
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3920939"
 /issue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 72"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; cloned unidirectionally. Primer: Oligo dt.
 Average insert size 2 kb. Library constructed by Life
 Technologies."

ORIGIN

Alignment Scores:

Pred. No.:	2,83e-103	Length:	899
Score:	1091.00	Matches:	210
Percent Similarity:	99.53%	Conservative:	0
Best Local Similarity:	99.53%	Mismatches:	0
Query Match:	49.46%	Indels:	1
DB:	2	Gaps:	0

US-10-798-532-4 (1-419) x BE895119 (1-899)

```

OY      113  LeuLeuaspValSerValSerGluLeuLeuTySerSerHisGlnGlyCySerMet 132
DB      1  CTCCTGGATGTCAGTGTTCGGAAATGCTTATATTCAGTCCAGGAGGTTGTCAG 60
OY      133  TTPMetILGlnHisCyAlaArgaspValLeuGlnAlaLeuAlaPheLeuHisGlu 152
DB      61  TGGATATACAGCATTTGTGCCAGAGATGTTTGGAGGCCCTTGTCTTCATCATGAG 120
OY      153  GATYTYValHisAlaAspLeuLeuProArgAsnILEuTrpSerAlaGluAsnGly 172
DB      121  GGGTATGTCATCGGAGCTCAACCAAGTAATTTGGAGTGTCAGAGATGATATG 180
OY      173  PheLeuLeuLeuAspPheGlyLeuSerPheLeuGlyAsnGlnAspVallyTyTle 192
DB      181  TTTAACTCATTCATTGGACTTGGACTTCAAGAGGCAATCAGATGTAAGTATAT 240
OY      193  GlnThaArgGlyTyTrgAlaProGlnAlaGlnLeuGlnAsnCyLeuAlaGlnAla 212
DB      241  CAGACAGACGGGTATCGGCTCCAGAGAGCAAAATTCGCTTGGCCCGCTGGC 300
OY      213  LeuGlnSerAspThrGlnCyThrSerAlaValAspLeuTrpSerLeuGlyLeileu 232
DB      301  CTGCAGAGTGATCAAAATGTACTCAGCTGTGATCTGGAGCTTAGAAATCATTTA 360
OY      233  LeuGluMetPheSerGlyMetLeuLeuHisThrValArgSerGlnGlyTyrAla 252
DB      361  CTGGAAATGTTCTCAGGAATGAATGAACATACATCAGATCTCAGGAATGGAAGCA 420
OY      253  AsnSerSerAlaIleLeuAspHisIlePheAlaSerIleValValAsnAlaIle 272
DB      421  AACAGTTGTCATATTTATGATCATATTTTGGCAAGTAAGCAGTGGGAATGCGCAAT 480
OY      273  ProAlaTyHisLeuArgAspLeuIleLeuSerMetLeuHisAspAspProSerArg 292
DB      481  CCGAGCTATACCTTAAGAGACCTTATCAAAAGCATGTTCAATGATATCAACAGAGA 540
OY      293  IleProAlaGluMetAlaLeuCySerProPhePheSerIleProPheAlaProHis 312
DB      541  ATTCCTGTAAGAAATGCGATTTGCGAGCCCATTCCTTATTCATTCCTTTGCCCC 600
OY      313  GluAspLeuValMetLeu-ProthrProVal 322
DB      601  GAAGATCTGTCATGCTTCCCACTCAGTT 631

```

RESULT 13

BI562420 772 bp mRNA linear EST 05-SEP-2001
 LOCUS BI562420 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5298674 5',
 DEFINITION mRNA sequence.

ACCESSION

BI562420
 VERSION BI562420.1 GI:15449747
 KEYWORDS EST.

SOURCE

ORGANISM Homo sapiens (human)

REFERENCE

1 (bases 1 to 772)
 NIH-MGC http://mgs.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgapbs-rt@mail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov

plate: LLM11755 row: p column: 03
 High quality sequence stop: 727.

FEATURES

source

1. .772
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5298674"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_97"
 /note="Organ: testis; Vector: pBluescriptR (modified
 pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
 (gtcgag); Oligo-dT primed using primer
 5'-TTTTTTTTTTTTTTVN-3', size-selected for average
 insert size 2.2 kb and normalized to R0T 5. This is a
 primary library enriched for full-length clones and
 constructed using the Cap-trapper method (Carninci, in
 preparation). Library constructed by M. Brownstein
 (NIH/NHGRI, National Institutes of Health). Note: this is
 a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	4,06e-102	Length:	772
Score:	1079.00 <td>Matches:</td> <td>212 </td>	Matches:	212
Percent Similarity:	96.82% <td>Conservative:</td> <td>1 </td>	Conservative:	1
Best Local Similarity:	96.35% <td>Mismatches:</td> <td>5 </td>	Mismatches:	5
Query Match:	48.91% <td>Indels:</td> <td>3 </td>	Indels:	3
DB:	4 <td>Gaps:</td> <td>1 </td>	Gaps:	1

US-10-798-532-4 (1-419) x BI562420 (1-772)

```

OY      92  LeuTYrGlyValPheThrIleHisPheSerProAsnValProSerArgCySerLeuLeu 111
DB      108  TTGTATGAGGTATTAACAATCACTTTCTCCAAATGTCATACAGCTGTGTGCTT 167
OY      112  GluLeuLeuaspValSerValSerGluLeuLeuTySerSerHisGlnGlyCySer 131
DB      168  GAATCTCTGATGTCAGTGTTCGGAAATGCTTATATTCAGTCCAGCANGTTGTTCC 227
OY      132  MetTrpMetILGlnHisCyAlaArgaspValLeuGlnAlaLeuAlaPheLeuHis 151
DB      228  ATGTGATATACAGCATTTGTGCCAGAGATGTTTGGAGGCCCTTGTCTTCATCAT 287
OY      152  GluGlyTYrValHisAlaAspLeuLeuProArgAsnILEuTrpSerAlaGluAsn 171
DB      288  GAGGCTATGTCATCGGAGCTCAACCAAGTAATTTGGAGTGTCAGAGATGA 347
OY      172  CyPheLeuLeuLeuAspPheGlyLeuSerPheLeuGlyAsnGlnAspVallyTyT 191
DB      348  TGTTTAACTCATTCATTGGACTTGGACTTCAAGAGGCAATCAGATGTAAGTAT 407

```

QY 192 IlegInThAspGlyTyrArgAlaProGluAlaGluLeuGlnAsnCysLeuAlaGlnA 211
 DB 408 ATTCAAGACAGACGGGATCGGGCTCCAGAGCAAGATTGCCAAATGCTTGCCCGCAGCT 467
 QY 212 GtyleuGlnSerAspThrGluCysThrSerAlaValAspLeuTrpSerLeuGlyIleIle 231
 DB 468 GGCCTGCAGAGATGATACAAATGTACTCAGCTGTGTGATCTGTGGAGCCTAGGAATCATTT 527
 QY 232 LeuLeuGluMetPheSerGlyMetLeuLysHisThrValArgSerGlnGluTrpLys 251
 DB 528 TTACTGGAAATGTTCTTCAGGAATGAAATGAAACATACAGTCAGATCTCAGGAATGAAAG 587
 QY 252 AlaSerSerSerAlaIleIleAspHisIle-PheAlaSerLysAlaValAlaAsnAlaI 271
 DB 588 GCACACAGTTCTGCTATTATTTGATCATTTATTTGCCAGTAAGCAGTGCTGATCCGCC 647
 QY 271 AlieProAlaTyrHisLeuArgAspLeuIleLysSerMetLeuHisAspAspProSerAr 291
 DB 648 AATTCCAGCCTATACACTTAAGAGACCTTATCAAAACATGCTTCATGATGATCAAGCAG 707
 QY 291 gArGlieProAlaGluMetAlaLeuGlySerProPhePheSerIleProPheAlaPro 310
 DB 708 AAGAATTCCTGCTGAATGACATTTGACAG-CCATTCTTAGCA--TCCTTTGGCCCT 761
 RESULT 14
 BU205357 794 bp mRNA linear EST 25-NOV-2002
 LOCUS 603104569F1 CSEQCHN03 Gallus gallus cDNA clone CHEST41j18 5', mRNA
 DEFINITION sequence.
 ACCESSION BU205357
 VERSION BU205357.1 GI:25370848
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 1 (bases 1 to 794)
 Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Boesch, E.,
 Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.,
 A Comprehensive Collection of Chicken CDNAS
 Curr. Biol. 12 (22), 1965-1969 (2002)
 2235534
 12445392
 COMMENT Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology
 (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.
 Location/Qualifiers
 1..794
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="White Leghorn, HiseX"
 /db_xref="taxon:9031"
 /clone="CHEST41j18"
 /tissue_type="whole embryo"
 /dev_stage="20-21"
 /lab_host="DH10B"
 /clone_1fb="CSEQCHN03"
 /note="Organ: whole embryo; Vector: pBluescript II KS(+);
 Site 1: EcoRI; Site 2: NotI; This normalized library was
 constructed from 1 million independent clones. cDNA
 synthesis was initiated using an oligo(dT) primer, using
 methylated C in the first strand synthesis reaction.
 Following this first strand reaction, double-stranded cDNA
 was blunt-ended, ligated to NotI adapters, digested with
 EcoRI, size-selected, and cloned into the NotI and EcoRI
 compatible sites of a custom modified MCS of the
 pBluescript (KS+) vector. The library was normalized in 2

rounds using conditions adapted from Soares et al., PNAS
 (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
 (1996): 791, except that a significantly longer
 reannealing hybridization was used."

ORIGIN

Alignment Scores:

Pred. No.:	1,11e-101	Length:	794
Score:	1075.00	Matches:	221
Percent Similarity:	89.10%	Conservative:	16
Best Local Similarity:	83.08%	Mismatches:	27
Query Match:	48.73%	Indels:	0
DB:	5	Gaps:	4

US-10-798-532-4 (1-419) x BU205357 (1-794)

QY 82 GtInLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPheSer 101
 DB 1 CAGTCCCGGCGGACCGCAACATCGTACTGTACGGCGGTGTCACCAACCATACTCG 60
 QY 102 ProAsnValProSerArgCysLeuLeuLeuGluLeuLeuValSerValSerLeuLeu 121
 DB 61 GCCAACGGCCCGTCCGCTGCTGCTGAGCTGTGATATACGGCTGTACAGCTG 120
 QY 122 LeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArgAsp 141
 DB 121 CTGCTGCACTCCAGCAACACAGGTCTGCTCCATGATGATGATCCAGCATCGCCCGCCAGC 180
 QY 142 ValLeuGluAlaLeuAlaPheLeuHisHisGlnGlyTyrValHisAlaAspLeuLysPro 161
 DB 181 GTCTGGAAGCCCTGGCTGCTCTGACACCAAAAGGCTACGTACAGCAGCAGCCTGAAGCCA 240
 QY 162 ArgAsnIleLeuTrpSerIleArgLeuGluGluCysPheLysLeuIleAspPheGlyLeuSer 181
 DB 241 CCGCAACATCTGTGAGCGCGGAGGAGAGAGTCTTAAAGCTCATTTGACTTTGAGACTTTTAC 300
 QY 182 Phe-LysGlnGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaProG1 201
 DB 301 TTCAAAAGGGGGAATTCAGAGATGTAATATTTTAAACAGACGGGTATCGGGCTCCAGA 360
 QY 201 uAlaGluLeuGlnAsnCysValuGlnAlaGlyLeuGlnSerAspThrGluCysThrSe 221
 DB 361 GGCAGAGCTGCAAGACTGCTGACAGGACAGGAGGCTCCAGAGTACAGAGATGACTTC 420
 QY 221 rAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLysLe 241
 DB 421 TCTGTGATCTGTGTGAGCTGGGAATGTGTTTACTGAAATGTTCTCAGGAATGAAACT 480
 QY 241 uLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHisI 261
 DB 481 GAAACATACAGTCCATATCCAGGAATGAAAGCAACAGTTTGCATCATGATGATCGCAT 540
 QY 261 ePheAlaSerLysAlaValAlaAsnAlaAlaIleProAlaTyrHisLeuArgAspLeuI 281
 DB 541 TTTTCCACTGAAGGGGTGTTAATCCAGCCATTCCAGTTATACCTCAGAGACTTAT 600
 QY 281 eLysSerMetLeuHisAspAsp-ProSerArgArgIleProAlaGluMetAlaLeuCysS 301
 DB 601 TAAAGCATGCTTCAATTGTGACCAAGCAAAAGCAGCTCTGTGAAAAGCTTATATGCA 660
 QY 301 eProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThrP 321
 DB 661 GCCCATCTTTCAGCAATCCCTTGCTCC-CATATTTGAAGATTGTGTGATGCTCCCAAG-C 718
 QY 321 rValLeuLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyLysGlnGluGluTyrG 341
 DB 719 CTGTCTGAGCGCTGCTGAATGTTCTTAAGCCAGAGGTCTCTGCAAGTGAAGAAAATACG 778
 QY 341 lAspValValGlu 345
 DB 779 AGAATATCTGGGAA 792
 RESULT 15

CF593363 834 bp mRNA linear EST 26-SEP-2003
LOCUS AGENCOURT15622279 NIH_MGC.147 Homo sapiens cDNA clone
DEFINITION IMAGE:10531268 5', mRNA sequence.
ACCESSION CF593363
VERSION CF593363.1 GI:36346826
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 834)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgsabbs-r@mail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: NDM621 row: m column: 05
High quality sequence scop: 660.
Location/Qualifiers
1. 834
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30531268"
/isue_type="Human Placenta"
/lab_host="DH10B TONa"
/clone_lib="NIH_MGC 147"
/note="Organ: Placenta; Vector: pBluescript; Site 1:
all-XhoI; Site 2: BamI; oligo-dT primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.3 kb and normalized to R07 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NHGRI/NHGRI, National Institutes of Health). Note: This is
a NIH_MGC library."

FEATURES

source
1. 834
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30531268"
/isue_type="Human Placenta"
/lab_host="DH10B TONa"
/clone_lib="NIH_MGC 147"
/note="Organ: Placenta; Vector: pBluescript; Site 1:
all-XhoI; Site 2: BamI; oligo-dT primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.3 kb and normalized to R07 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NHGRI/NHGRI, National Institutes of Health). Note: This is
a NIH_MGC library."

ORIGIN

Alignment Scores:

Pred. No.: 3.54e-101 Length: 834
1070.50 Matches: 213
Percent Similarity: 95.13% Conservative: 2
Best Local Similarity: 94.25% Mismatches: 5
Query Match: 48.53% Indels: 6
DB: 7 Gaps: 1

US-10-798-532-4 (1-419) x CF593363 (1-834)

QY 1 MetAAGlySerGlyCySAATpGlyAAGlnProProArgPheLeuGlnAaPheGly 20
DB 155 ATGGCGGATCCGGCTGGCGTGGGCGGAGCCGCCGCTTCTGGAGGCTTCGGG 214
QY 21 ArgLeuTrpGlnValInserArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
DB 215 CGGCTGTGGCAGGTACAGAGCCGTCTGGTGGGCTCCCTCCGCTGGGTATCGGGTT 274
QY 41 ArgCyAGlyAanProGlySerProGlyAlaLeuArgGlnPheLeuProProGly 60
DB 275 CGCTGTGGGCGAACCCTGGCTGGCCCGGCGCCCTCAAGCACTTCTTGGCCGCAAGA 334
QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgGlyGlnArgAlaLeu 80

DB 335 ACCACGGGGCTGGCGCTCTGCCCGCAGATGCTTCCGAAAGAGAGGGCGCTG 394
QY 81 GlnGlnLeuGlnGlyValArgAanIleValThrLeuTyrGlyValPheThrIleHisPhe 100
DB 335 GAAAGTTCGAGGGGTCAAGAAACATCGTACTTTGTATGAGAGTCTTACATCCACTT 454
QY 101 SerProAanValProSerArgCyLeuLeuLeuGlnLeuLeuAanPValSerValSerGlu 120
DB 455 TCTCCAAATGTCCATCAAGCTGTCTGTGGCTTGAATCTCGAGATGATGATGATGATG 514
QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCySerMetTrpMetIleGlnHisCySAAlaArg 140
DB 515 TTGCTCTTATATTCAGTCACAGGAGTGTTCATGATGATGATGATGATGATGATGATG 574
QY 141 AapValLeuGlnAlaLeuAlaPheLeuHisHisGlnGlyTyrValHisAlaAanPheLeu 160
DB 575 GATGTTTGGAGGCCCTTCTTTCTTCATCATGAGGCTATGTCATCGGACCTCAA 634
QY 161 ProArgAanIleLeuTrpSerAlaGlnAanGlyCyPheLeuLeuIleAanPheGly-Le 180
DB 635 CCACGTAACATATGTGGAGTGCAGAGATGAATGTTTANACTATTCATTCGACTTGGACTT 694
QY 180 userPheLeuGlnGlyAanGlnAanPValLysTyrIleGlnThrAanPValTyrArg-AlaP 200
DB 695 AGCTTCAAGAAAGCAATCAGAGATGTAAATTCAGACAGCGGGTATCGGGCTC 754
QY 200 roGlnAlaGlnLeuGlnIn-AanCySAAlaA--GlnAlaGly-LeuGlnInserAanPThrG 218
DB 755 CAGAGCAGAAATTCGAAATAATGCTTGGGCCCCAGGCTTGGCCCTGCAGAGATGATCAGA 814
QY 218 u-CysThr 220
DB 815 AATGTACC 822

Search completed: November 29, 2004, 13:27:52
Job time : 3168.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 29, 2004, 11:25:02 / Search time 502 Seconds
(without alignments)
4381.488 Million cell updates/sec

Title: US-10-798-532-4

Perfect score: 2206
Sequence: 1 MAGSGCAGABPPRFLEAFG.....VATFYPPLSAVKRGYLYQTLL 419

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=alp
-Q/cgq2.1/USPTO.spool_p/US10798532/runat_29112004.112452.15882/apd_query.fasta.1.1166
-DB=N Geneseq 23Sep04 -QPM=fastap -SUPFIX=ring -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITs=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human4.0.cdi
-LIST=45 -DOCALL=200 -THR SCORE=500 -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTMT=ptc -NORMext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10798532.QCGN_1.1.886.@runat_29112004.112452.15882 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: N_Geneseq_23Sep04:*
2: geneseqn1808:*
3: geneseqn1908:*
4: geneseqn2008:*
5: geneseqn2018a:*
6: geneseqn2018b:*
7: geneseqn2028a:*
8: geneseqn2028b:*
9: geneseqn2038a:*
10: geneseqn2038b:*
11: geneseqn2038c:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARYS

Result No.	Score	Query Match	Length	DB ID	Description
1	2206	100.0	1260	3	AAZ51356 Human tra
2	2203	99.9	1260	3	AAZ51355 Human wil
3	2195	99.5	1260	4	AAI6829 Human pro
4	2195	99.5	2008	6	ABQ9383 Human cod
5	2195	99.5	2622	4	AAI6828 Human pro
6	2195	99.5	4065	4	AAH72714 Human cer

7	2195	99.5	4065	5	ADL45238 Human ova
8	2188	99.2	1260	4	AA606734 Polynucle
9	2188	99.2	1824	4	AAI18818 Human kin
10	2183	99.0	3244	10	AB52857 Primary r
11	2042.5	92.6	3598	12	ADQ24401
12	789	35.8	529	4	AAH70764 Human cer
13	766.5	34.7	722	3	AAH02535 Human col
14	729	33.0	470	4	AAH71966 Human cer
15	636	28.8	461	4	AAH69101 Human cer
16	559.5	25.4	411	6	ABQ99225 Human ORF
17	361	16.4	231	3	AAH01259 Human col
18	311.5	14.1	419	5	AD173740 Human ova
19	311.5	14.1	419	5	AD167345 Human ova
20	284	12.9	546	6	ABQ53127 Oligonuc1
21	284	12.9	546	6	ABQ53126 Oligonuc1
22	284	12.9	547	6	ABQ44465 Oligonuc1
23	284	12.9	547	6	ABQ44464 Oligonuc1
24	255	11.6	546	6	ABQ53129 Oligonuc1
25	255	11.6	546	6	ABQ53128 Oligonuc1
26	255	11.6	547	6	ABQ44463 Oligonuc1
27	255	11.6	547	6	ABQ44462 Oligonuc1
28	250	11.3	442	9	ACH28557 Human adu
29	249	11.3	1065	10	ADD21399 Golden ha
30	246.5	11.2	1400	10	ADD21407 Rat cycl1
31	245	11.1	1708	10	ADD21395 Mouse cyc
32	240.5	10.9	2472	3	AA45716 Arabidops
33	240	10.9	1297	10	ADD21393 Human cyc
34	240	10.9	2213	6	ABR84073 Human cDN
35	240	10.9	2213	10	ADD21403 Chinese h
36	240	10.9	2213	11	AD131711 Human cDN
37	240	10.9	2213	12	ADH17138 Human cdc
38	240	10.9	2328	10	ACC72837 Human can
39	240	10.9	2734	12	ADQ25034 Human sof
40	239	10.8	1650	12	AD145326 Rice 18op
41	237	10.7	1209	10	ADD21401 Golden ha
42	236.5	10.7	1240	6	AB188413 Pain regu
43	236.5	10.7	1335	2	AAV71073 CDK2-gree
44	236.5	10.7	1655	3	AAA0869 Rat stres
45	236	10.7	1297	6	ABQ60787 Human 10C

ALIGNMENTS

RESULT 1
ID: AAZ51356 Standard; DNA; 1260 BP.

AC: AAZ51356;

DT: 06-JUN-2000 (first entry)

DE: Human transdominant mutant serine/threonine kinase KIS (hKIS) gene.

XX 20-AUG-1999; 99WO-US018903.
XX
XX 21-AUG-1998; 98US-0097710P.
XX
XX (NABEL/) NABEL G J.
XX (NABEL/) NABEL E G.
XX
XX Nabel GJ, Nabel EG;
XX
XX WPI: 2000-237648/20.
XX P-PSDB; AAY70306.
XX
XX Novel serine/threonine kinase hKIS polynucleotides and polypeptides used
XX for inhibiting the cyclin kinase inhibitor p27, and so alter cell
XX proliferation.
XX
XX Claim 7; Page 59; 70pp; English.
XX
XX The present DNA sequence encodes a transdominant mutant human KIS (hKIS),
XX constructed by site directed mutagenesis. A single nucleotide
XX substitution (A to G) results in a lysine to arginine change in the
XX protein sequence. hKIS is a serine/threonine kinase that acts as an
XX inhibitory kinase of cyclin-dependent kinase inhibitor (CKI), p27. hKIS
XX controls cell proliferation and is localised predominantly in the
XX nucleus. It binds to CKI p27 and inhibits its ability to arrest cells in
XX G1 phase. The hKIS sequences are used to modulate cell proliferation and
XX treat cell proliferative and vascular diseases. The polynucleotide
XX sequence may be used in gene therapy to treat vascular disorders such as
XX restenosis or atherosclerosis
XX

XX Sequence 1260 BP; 310 A; 280 C; 336 G; 334 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1 44e-207	Length:	1260
Score:	2206.00	Matches:	419
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	3	Gaps:	0

US-10-798-532-4 (1-419) x AAZ51356 (1-1260)

QY	1	MetAlGlySerGlyCyAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly	20
DB	1	ATGGCGGAGATCCGGCTGCGCGCTGGGGCGCGGACCGCGGCTTTCGAGGCGCTTCGGG	60
QY	21	ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal	40
DB	61	CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCTCCGCGGTGATCGGGTT	120
QY	41	ArgCyAspGlyValAspProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly	60
DB	121	CGCTGCTGGCGCAACCTGCTGCGCCCGCGGCGCTCAAGGACATTTCTTGGCCGCAAGA	180
QY	61	ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgGlyGluArgAlaAlaLeu	80
DB	181	ACCAACGGGGGCTCGGCGCTCTGCGCGCGAGTATGTTTCGCAAGAGAGGGCGGCTG	240
QY	81	GluGlnLeuGlnGlnValHisArgGlnIleValThrLeuTyrGlyValaPheThrIleHisPhe	100
DB	241	GAACAGTTGGAGGTACAGAAACATCGTACTTTGTATGAGAGTGTTCACATCCACTTT	300
QY	101	SerProAsnValProSerArgCySLeuLeuLeuGluLeuLeuAspValSerValSerGlu	120
DB	301	TCTCCAAATGTGCATCACCGCTCTGCTTGTGAACTCCTGATGTCACTGTTTCGGA	360
QY	121	LeuLeuLeuTyrSerSerHisGlnGlyCySerMetTrpMetIleGlnHisCyAlaArg	140
DB	361	TTCGCTTATATTCAGTCAACAAGGGTGTTCATGTGATGATACAGATTGCGCCGA	420
QY	141	AspValLeuGluAlaLeuAlaPheLeuHisHisGlnGlyTyrValHisAlaAspLeuVal	160

DB	421	GATGTTTGGAGGCCCTTCTTTCATCATGAGGCGTATGTCATGCGGACCTCAAA	480
QY	161	ProArgAsnIleLeuTrpSerAlaGluAsnGluCySPhelyLeuIleAspPheGlyLeu	180
DB	481	CCACGTACATATTTGTGAGTGCAGAGATGATGATTTTAACTCATGTGACTTGGACTT	540
QY	181	SerPheLeuGlnGlyValAsnGlnAspValIleValThrLeuTyrGlyValaPhe	200
DB	541	ACCTTCAAGAGGCAATCAGATGTAAAGTATATTCAGACAGAGGGATGCGGCTCA	600
QY	201	GluAlaGluLeuGlnAsnCySLeuAlaGlnIleValGlnGlnSerAspThrGlyCySer	220
DB	601	GAAGCAGAAATTCGAAATTTGCTTGGCCAGCGCTGGCGAGTGTATACAGATGATC	660
QY	221	SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnMetPheSerGlyMet	240
DB	661	TGAGCTGTTGATCTGTGGAGCCTTAGAATCATTTTACGAAATGTTCTCAGGAATGAA	720
QY	241	LeuValHisThrValArgSerGlnGluTyrPlyAlaAsnSerSerAlaIleIleAspHis	260
DB	721	CTGAAACATACAGTCAAGATCTCAGAAATGAAAGCAAGTTCGTATTAATTGATCAC	780
QY	261	IlePheAlaSerIleAlaValAlaAlaIleProAlaTyrHisLeuArgAspLeu	280
DB	781	ATATTTGCCAGTAAAGCAGTGTGATGCCGCAATTCAGCTTACACTTACAGACCTT	840
QY	281	IleIleSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeu	300
DB	841	ATCAAAAGCATCTTCATGATGATCCAGCAGAAAGATTCCTCGTGAATGCAATTGGC	900
QY	301	SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuPro	320
DB	901	AGCCCATTTCTTGAATTCCTTTTCCCTCATTTGAAGATCTGTCAATGCTTCCACT	960
QY	321	ProValIleuArgLeuLeuAsnValIleuAspAspArgTyrLeuGlyAsnGluGluTyr	340
DB	961	CCAGTGTCAAGCTGTGATGTGCTGAGTGTGATTTCTTGGAAATGAAAGAAATNT	1020
QY	341	GluAspValAlaGluAspValIleGlnGluGlyCySglnGlyTyrGlyProValIle	360
DB	1021	GAAAGATTTGTGAAGATCTTAAGAGGAGGTCTCAAAATATGACCACTGTATCTCTA	1080
QY	361	LeuValProIleGluAsnProGlyValArgGlyGlnValPheValGluTyrAlaAsn	380
DB	1081	CTTGTTCCCAAGAAATCTCTGCAAGAGCAAGTCTTGTGATGATGCAAAATCTGGT	1140
QY	381	AspSerIleValAlaGlnIleLeuLeuTyrGlyValArgMetPheAspGlyVal	400
DB	1141	GATTCGAAAGCTGCCGAGAAATTAATGACTGAGAGAGATTTGATGGAGATTTGTG	1200
QY	401	AlaThrPheTyrProLeuSerAlaTyrIleValArgGlyTyrLeuTyrGlnThrLeu	419
DB	1201	GCTACATTTACCCGCTGAGTGTCTTCAAGAGGGATATCTGTATCAAACTTGTCTT	1257

RESULT 2
AAZ51355
ID AAZ51355 standard; DNA; 1260 BP.
XX
AC AAZ51355;
XX
DT 06-JUN-2000 (first entry)
XX
DE Human wild type serine/threonine kinase KIS (hKIS) gene.
XX
KW KIS: human; hKIS: serine/threonine kinase; cell proliferation; G1 phase;
KW dominant gene; cyclin-dependent kinase inhibitor; CKI: p27; modulator;
KW treatment; cell proliferative disease; vascular disorder; gene therapy;
KW atherosclerosis; restenosis; ds.
XX
OS Homo sapiens.
XX
FT Key Location/Qualifiers
CDS 1..1260


```

FT      /-tag= a
FT      /product= "Human wild type KIS protein"
FT      /transl_except= (pos:556..561, Asn)
XX
XX      MO200011165-A1.
XX
XX      02-MAR-2000.
XX
XX      20-AUG-1999; 99MO-US018903.
XX
XX      21-AUG-1998; 98US-0097710P.
XX
XX      (NABE/) NABEL G J.
XX      (NABE/) NABEL E G.
XX
XX      Nabel GJ, Nabel EG;
XX
XX      WPI; 2000-237648/20.
XX      P-PSDB; AA970305.
XX
XX      Novel serine/threonine kinase hKIS polymucleotides and polypeptides used
XX      for inhibiting the cyclin kinase inhibitor p27, and so alter cell
XX      proliferation.
XX
XX      Claim 1; Page 56; 70pp; English.
XX
XX      The present DNA sequence encodes a wild type human KIS (hKIS), isolated
XX      from a human B-cell library, using a yeast two hybrid screening system.
XX      hKIS is a serine/threonine kinase, that acts as an inhibitory kinase of
XX      cyclin-dependent kinase inhibitor (CKI), p27. hKIS is a dominant gene
XX      that controls cell proliferation and is localised predominantly in the
XX      nucleus. It binds to CKI p27 and inhibits its ability to arrest cells in
XX      G1 phase. It has 99% homology to rat serine/threonine protein kinase KIS.
XX      The hKIS sequences are used to modulate cell proliferation and treat cell
XX      proliferative and vascular diseases. The polymucleotide sequence may be
XX      used in gene therapy to treat vascular disorders such as restenosis or
XX      atherosclerosis
XX
XX      Sequence 1260 BP; 311 A; 280 C; 335 G; 334 T; 0 U; 0 Other;
XX
XX      Alignment Scores:
XX      Pred. No.: 2,84e-207 Length: 1260
XX      Score: 2203.00 Matches: 418
XX      Percent Similarity: 100.00% Conservative: 1
XX      Best Local Similarity: 99.76% Mismatches: 0
XX      Query Match: 99.86% Indels: 0
XX      DB: 3 Gaps: 0
XX
XX      US-10-798-532-4 (1-419) x AA251355 (1-1260)
XX
QY      1 MetAlaGlySerGlyCysAlaATTPGlyAlaGluProProArgPheLeuGluAlaPheGly 20
DB      1 ATGGCCGGATCCCGGCTGCGCGCGCGCGAGCCCGCGGTTTTCGAGGGCTTCGG 60
QY      21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTYrArgVal 40
DB      61 CGGCTGTGGCAGGTACAGACCCCTCTGCGTACGGGCTCTCCCGCTCGGTATCCGGTT 120
QY      41 ArgCysCysGlyValAsnProGlySerProProGlyValAlaLeuArgGlnPheLeuProProGly 60
DB      121 CGCTGCTGGCGCAACCTGGCTGCGCCCGCGCGCCCTCAAGAGTTCTTGGCGCAGGA 180
QY      61 ThrThrGlyValAlaAlaSerAlaAlaGluTYrGlyPheArgGlyGluArgAlaAlaLeu 80
DB      181 ACCACCGGGGCGGCGCTCTGCGCGCGAGTATGTTTCCGAAAGAGGGCGGCGCTG 240
QY      81 GluGlnLeuGlnGlnHisArgAsnIleValThrLeuTYrGlyValPheThrIleHisPhe 100
DB      241 GAACAGTTGGAGGCTCAGAGAAACATCGTCACTTTGATGAGATGTTTACATCCACTTT 300
QY      101 SerProAsnValProSerArgCysLeuLeuLeuGlnLeuLeuAspValSerValSergIu 120
DB      301 TTCCAAAATGTGCATCAGCGCTGTCTGTGCTTGAACCTCGATGATGCAGTGTTTCGAA 360

```

```

QY      121 LeuLeuLeuTYrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
DB      361 TTGCTCTTATATTCAGTACACAGGGTTGTTTCAATGGATGATACAGATTGGCCCGCA 420
QY      141 AspValLeuGlnAlaLeuAlaPheLeuHisHisGluGlyTYrValHisAlaAspLeuLys 160
DB      421 GATGTTTGGAGGCCCTTCTTCTTCATCATGAGGCTATGTCCATGCCGACCTCAAA 480
QY      161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLeuIleAspPheGlyLeu 180
DB      481 CCACGTAACATATTTGAGTGCAGAGATGATGATTTAACTCAATTGACTTTGACTT 540
QY      181 SerPheLysGluGlyValAsnGlnAspValLysTYrIleGlnThrAspGlyTYrArgAlaPro 200
DB      541 AGCTTCAAAAGAGGCAATCAGATGTAAAGTATATTCAGACAGAGGGATGGGCTCCA 600
QY      201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGlyCysThr 220
DB      601 GAAGCAGAAATTGCAAAATTGCTTGGCCCGCGCTGGCTGCAGAGTATCAGAAATGTA 660
QY      221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
DB      661 TCAGCTGTGATCTGTGAGCTCAGAAATCATTTTATCTGAAATGTCTCAGAAATGAAA 720
QY      241 LeuLysHisThrValArgSerGlnGluTYrPlyAlaAsnSerSerAlaIleIleAspHis 260
DB      721 CTGAAACATACGATGATCTCAGAAATGAAAGCAACAGTTCGCTTATTTATATCAC 780
QY      261 IlePheAlaSerLysAlaValAlaAsnAlaAlaIleProAlaTYrHisIleuArgAspLeu 280
DB      781 ATATTTCAGATGAAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY      281 IleLysSerMetLeuHisAspAspProSerArgAlaGlyIleProAlaGluMetAlaLeuCys 300
DB      841 ATCAAAAGATCTTCATGATATATCAAGCAGAAATATTCGTGAATGACATTGTGC 900
QY      301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
DB      901 AGCCCATTTTATGATCTCTTTGCTTCCCTCATATTTGAAGATCTGTCATGCTTCCACT 960
QY      321 ProValLeuArgLeuLeuAsnValLeuAspAspTYrLeuGlyValAsnGluGluTYr 340
DB      961 CCAGTGTCTAAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY      341 GluAspValValGluAspValLysGluGluCysGlnLysTYrGlyProValValSerLeu 360
DB      1021 GAAGATGTTGTGAAGATGTAAAGAGAGTCTCAAAATATGACCACTGATCTCTTA 1080
QY      361 LeuValProLysGluAsnProGlyValArgGlyGlnValPheValGluTYrAlaAsnAlaGly 380
DB      1081 CTGTGTCAAAGAAATCTCTGCAAGAGCAAGTCTTGTGATGATGCAAAATGCTGCT 1140
QY      381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
DB      1141 GATTCAAAGCTGCCAGAAATTAATGACTGGAAGAGATTTGATGGGAAGTTTGTGTGG 1200
QY      401 AlaThrPheTYrProLeuSerAlaTYrLysArgGlyTYrLeuTYrGlnThrLeuLeu 419
DB      1201 GCTACATTTCTACCCGCTGATGCTTCAAGAGGGATATCTGATCAAACTTGTCTT 1257

```

RESULT 3
AA166829 standard; cDNA, 1260 BP.
ID AA166829;
AC AA166829;
XX
XX 07-JAN-2002 (first entry)
DE Human protein kinase polypeptide 13887 coding sequence.
XX
XX Protein kinase; 3714; 16742; 23546; 13887; cancer; bone disorder; human;
KW cytostatic; antiinflammatory; immunosuppressive; cardiac; hepatotropic;

KM gene therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1260
FT /tag= a
FT /product= "protein kinase 13887"
XX
XX MO200173050-A2.
XX
XX 04-OCT-2001.
XX
XX 23-MAR-2001; 2001WO-US009483.
XX
XX 24-MAR-2000; 2000US-0191846P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Meyers R;
XX
XX WPI: 2001-611632/70.
XX
XX P-PSDB; AAG65767.
XX
XX
XX New human protein kinase polypeptides, 3714, 16742, 23546 and 13887,
PT useful in diagnosis of cancer or cellular proliferation or
PT differentiation disorders and to screen for polypeptide modulators useful
PT to treat such conditions.
XX
XX
XX Claim 1; F1g 10A-C; 169pp; English.
XX
XX The invention provides novel human protein kinase polypeptides, 3714,
CC 16742, 23546 and 13887 and nucleic acid molecules encoding them. The
CC protein kinase polypeptides can be expressed by standard recombinant
CC methodology. 3714, 16742, 23546 or 13887 nucleic acids and polypeptides
CC are useful for diagnostic and screening methods to identify subjects (at
CC risk of) having cancer or cellular proliferation and/or differentiation
CC disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and
CC modulators are useful for the treatment of cancer, particularly colon
CC cancer or cellular proliferation and/or differentiation disorders. Other
CC disorders associated with 3714, 16742, 23546 or 13887 expression or
CC activity that can be treated include bone related disorders, inflammatory
CC disorders, autoimmune diseases, cardiovascular disorders and liver
CC diseases. The present sequence represents a human protein kinase
CC polypeptide 13887 coding sequence
XX
XX Sequence 1260 BP; 312 A; 279 C; 334 G; 335 T; 0 U; 0 Other;
SO
Alignment Scores:
Pred. No.: 1,75e-206 Length: 1260
Score: 2195.00 Matches: 417
Percent Similarity: 99.76% Conservative: 1
Best Local Similarity: 99.52% Mismatches: 1
Query Match: 99.50% Indels: 0
DB: 4 Gaps: 0
US-10-798-532-4 (1-419) x AA166829 (1-1260)
QY 1 MetAlAGlySerGlyCysAlaATpGlyAlaGluProProArgPheLeuGluAlaIlePheGly 20
Db 1 ATGACCGGAAATCCGAGCTGCGCTGGGGCGCGAGCCCGCGCTTTCTGGAGGGCTTCGGG 60
QY 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTYrArgVal 40
Db 61 CGGCTGTGGAGGTAAGAAGCCGCTGTGGTAGCGGCTCTCCGCTCGGTATCGGGTT 120
QY 41 ArgCysGlyGlnAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
Db 121 CGCTGCTGCGGCAACCTGCTGCGCCCGCGGCGCTCAAGAGATCTTTCGCGCAGGA 180
QY 61 ThrThrGlyValAlaAlaSerAlaAlaGluTYrGlyPheArgGlyGlnAlaGlyAlaLeu 80
Db 181 ACCACCGGGGCTGCGCTCTGCGCGAGATATGTTTCCGCAAAAGAGGGCGGCTG 240

QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTYrGlyValPheThrIleHisPhe 100
Db 241 GAACAGTTCGACGGCTCACAGAAACATCGTGACTTGTATGAGGTGTTAACATCCACTTT 300
QY 101 SerProAsnValProSerArgCysLeuLeuLeuGlnLeuLeuAspValSerValSerGlu 120
Db 301 TCTCCAAAGTGCATCAAGCTGTGTGTGCTTGAACTTCGGAATGTCAGTGTTCGGAA 360
QY 121 LeuLeuLeuTYrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
Db 361 TTGCTCTTATATTCAGTCACAGGGTGTTCATGTGATATACAGATTGTCCCGA 420
QY 141 AspValLeuGlnAlaLeuAlaPheLeuHisIleGlnGlyTYrValHisAlaAspLeuYs 160
Db 421 GATGTTTGAAGCCCTTGCTTTCTTCATCATGAGGGCTATGTCATGCGGACCTCAAA 480
QY 161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLeuIleAspPheGlyLeu 180
Db 481 CCACGTAAATATTTGGAGTGCAGAGATGATGTATTTAACTATGACTTTGACTT 540
QY 181 SerPheLeuGlyValAsnGlnAspValIleTYrIleGlnThrAspGlyTYrArgAlaPro 200
Db 541 AGCTTCAAAGAAAGCAATCAGATGTAAAGTATATTCAGACAGCGGATGCGCTCA 600
QY 201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
Db 601 GAAGCAGATTCCAAAATGCTTGCGCCGAGCTGCGCAGAGTATCAGATGATCAGAAATGACC 660
QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetCys 240
Db 661 TAGCGTTGATCTGTGGAGCCTAGAGAAATCTTATCTGGAATGTTCACAGAAATGAAA 720
QY 241 LeuYHisIleThrValArgSerGlnGluTrpIleAlaAsnSerSerAlaIleIleAspHis 260
Db 721 CTGAACATACATCAGTCAAGATCTCAGAAATGGAAGCAAGTTCGTATTTATATCAC 780
QY 261 IlePheAlaSerIleAlaValAlaIleAsnAlaIleProAlaTYrHisLeuArgAspLeu 280
Db 781 ATATTTCGACAGTAAAGCAGTGTGAATGCCGCAATTCACGCTTAACCTTAAGACACCTT 840
QY 281 IleYsSerMetLeuHisAspAspProSerArgAlaIleProAlaGluMetAlaLeuCys 300
Db 841 ATCAAAAGATCTTCATGATATATCCAAAGCAGAAATTCCTGCGAAATGGCATTTGTGC 900
QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
Db 901 AGCCCATTTTTCAGCATTCCTTTTCCCTCATATTTGAAGATCTGTCAATGCTTCCACT 960
QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTYrLeuGlyAsnGluGluTYr 340
Db 961 CCAGTCTCAAGACTGCTGAATGTGCTGATGATGATTTATCTTGAGAAATGAAAGCAATAT 1020
QY 341 GluAspValValGluAspValIleYsGluGluCysGlnIleTYrGlyProValIleSerLeu 360
Db 1021 GAAGATGTTGTGAAGATGTAAAGAGAGTGTCAAAATATAGACCAAGTGTATCTCTCA 1080
QY 361 LeuValProIleGlyAsnProGlyArgGlyGlnValPheValGluTYrAlaAsnAlaGly 380
Db 1081 CTTGTTCCAAAGAAATCTCTGCGCAGAGACAAAGTCTTTGTTGAGATACCAATGTGGT 1140
QY 381 AspSerIleAlaAlaGlnIleLeuLeuThrGlyArgMetPheAspGlyIlePheValVal 400
Db 1141 GATTCCAAAGCTGCGCAGAAATTAAGTCACTGAGAGATGTTGATGGGAAAGTTGTGTG 1200
QY 401 AlaThrPheTYrProLeuSerAlaTYrIleYsArgGlyTYrLeuLeuGlnThrLeuLeu 419
Db 1201 GCTACATTTCTACCCGCTGAGTGTCTACAGAGGGATATCTGTATCAAACTTGTCTT 1257
RESULT 4
ABQ99383
ID ABQ99383 standard; cDNA; 2008 BP.
XX

AC ABQ99383;
 XX 25-FEB-2003 (first entry)
 XX
 DE Human coding sequence SEQ ID 116.
 XX
 XX Human, expressed sequence tag; EST; haematopoietic disorder;
 KM central nervous system disease; viral infection;
 KM peripheral nervous system disease; non-healing wound; infectious disease;
 KM immune deficiency; immune disorder; bacterial infection; allergy; cancer;
 KM fungal infection; autoimmune disorder; coagulation disorder; neutrotropic;
 KM anti-allergic; anti-inflammatory; immunosuppressive; neuroprotective;
 KM cytostatic; haemostatic; virocidic; antibacterial; fungicide;
 KM immunostimulant; cerebroprotective; gene therapy; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200259260-A2.
 XX
 XX 01-AUG-2002.
 PD
 XX
 PF 16-NOV-2001; 2001MO-US042950.
 XX
 PR 17-NOV-2000; 2000US-00714936.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao Q;
 PI Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;
 PI
 DR N-PSDB; ABP64797.
 XX
 PT New isolated polynucleotide, useful in research, diagnostic or
 PT therapeutic methods, e.g. preventing or treating disorders involving
 PT aberrant protein expression or biological activity.
 XX
 PS Claim 1; SEQ ID NO 116; 394pp; English.
 XX
 XX The present invention relates to novel human coding sequences (ABQ99268-
 CC ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in
 CC diagnostic, diagnostic and research methods. The polynucleotides may be
 CC used in the field of molecular biology as hybridisation probes, primers
 CC for PCR, for chromosome and gene mapping, for the recombinant production
 CC of protein, or in generation of anti-sense DNA or RNA. The
 CC polynucleotides are useful in diagnostics as expressed sequence tags
 CC (ESTs) for identifying expressed genes or for physical mapping of the
 CC human genome. The proteins may be used as molecular weight markers, or as
 CC nutritional sources or supplements. The proteins may be used to maintain
 CC and expand cell population in a totipotentia or pluripotentia state
 CC useful for re-engineering damaged or diseased tissues, transplantation,
 CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The
 CC polynucleotides and proteins are useful for preventing, treating or
 CC ameliorating disorders involving aberrant protein expression or
 CC biological activity, e.g. haematopoietic disorders, central/peripheral
 CC nervous system diseases, mechanical and traumatic disorders, non-healing
 CC wounds, immune deficiencies and disorders, infectious diseases caused by
 CC viral, bacterial or fungal infection, autoimmune disorders, allergic
 CC reactions and conditions, coagulation disorders, or cancer. The
 CC polynucleotide sequences of the invention were assembled from ESTs
 CC isolated mainly by sequencing by hybridisation, and in some cases,
 CC sequences obtained from one or more public databases. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 2008 BP; 491 A; 451 C; 488 G; 578 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3,46e-206 Length: 2008
 Score: 2195.00 Matches: 417
 Percent Similarity: 99.76% Conservative: 1
 Best Local Similarity: 99.52% Mismatches: 1

Query Match: 99.50% Indels: 0
 DB: 6 Gaps: 0
 US-10-798-532-4 (1-419) x ABQ99383 (1-2008)
 QY 1 MetAlaGlySerGlyCysAlaATrPglYAlaGluProPArGpHeuGluAlaAheGly 20
 Db ATGCGCGGATCCGGCTGCGCTGGGGCGGAGCGCGCGGCTTTCTGAGAGCGCTTGGG 138
 QY 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerSerAlaSerAlaYTrArgVal 40
 Db CGGCTTGCGAGGTACAGAGCGCTTGGGTACCGGCTTCCGCTCGGTATCGGGTT 198
 QY 41 ArgCysCysGlyAspProGlySerProGlyAlaLeuArgGlnPheLeuProGly 60
 Db CGGCTGCTGGCGAACCTTGGCTGCGCCCGGCGCTTCAAGCATTTCTTGGCGCGAG 258
 QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTrpGlyPheArgLysGluArgAlaLeu 80
 Db ACCACCGGGGCTGCGGCTTGGCGCGAGTATGTTTCGCAAGAGAGGGCGGCTG 318
 QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTrpGlyValPheThrIleHisPhe 100
 Db GAACAGTTCAGGGTCCAGAAACATCGACTTGTATGAGTGTTCACATCCACTTT 378
 QY 101 SerProAsnValProSerArgCysLeuLeuGluLeuAspValSerValSerGlu 120
 Db 379 TCTCCAAATGTGCATCAGCTGTCTGTCTTGAATCTCGATGTCAGTGTTCGAA 438
 QY 121 LeuLeuLeuTrpSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
 Db 439 TTGCTCTTATATTCAGTCCAGCGGGTGTTCATGTGATGATACAGCATTTGGCCGA 498
 QY 141 AspValLeuGlnAlaLeuAlaPheLeuHisIleGluGlyTrpValHisAlaAspLeuLys 160
 Db 499 GATGTTTGAAGCCCTTCTTCTTCATCATGAGGGTATGTCATCGGACCTCAAA 558
 QY 161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
 Db 559 CACGTAACATTTGTGAGTGCAGAGTAAGTATTTAACTCATTTGACCTTGGACTT 618
 QY 181 SerPheLeuGluGlyAsnGlnAspValLysTrpIleGlnThrAspGlyTrpArgAlaPro 200
 Db 619 AGTTTAAAGAGGCAATCAGATGTAAAGTATTTACACAGACGGGTATGGGCTCCA 678
 QY 201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
 Db 679 GAAGCAGAAATTCGAAATTCCTTGGCCGAGCTGGCTGCAGAGTATCAGATGTACC 738
 QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
 Db 739 TCAGCTGTGATCTGGAGCTTGAAGATCAATTTTCTGAAATGTTCTCAGGAAGAA 798
 QY 241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
 Db 799 CTGAACATACAGTCACTCAGATTCAGAAATGGAAGCAAGTTCTGCTATTTGATCAC 858
 Db 859 ATATTTTCCAGTAAGCAATGCGATGCAATTCGCAATTCAGCTTATCACTTAAGACCTT 918
 QY 261 IlePheAlaSerLysAlaValAlaAsnAlaAlaIleProAlaTrpHisLeuArgAspLeu 280
 Db 281 IIElySerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
 Db 919 ATCAAAAGCATGCTTCATATATATCCACAGAGAAATTCCTGCGAATGGCATTTGTC 978
 QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
 Db 979 AGCCATTTCTTAAAGATTCCTTTTGGCCCTCATATTAAGATCTGGTCAATGCTTCCACT 1038
 QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspArgTrpLeuGlyAsnGluGluGlyTrp 340
 Db 1039 CCAGTCTCTAAGACTGCTGAATGTGCTGATGATGATATCTTGAAGAAAGAGAAATAT 1098

QY	341	GIUAAGVAlValGluLupBPValLLYVSGluGluCQVSGluLysPyrGlyPProValValSerLeu	360
Db	1099	GAAGAAGTGTAGAAAGATGTAAAAAGAGAGAGTGTCAAAAAATATGACCAAGTGTATCTTA	115
QY	361	LeuValProLysGluLupBPProGlyArgGlyGlnValPheValGluTyrValaenalagly	380
Db	1159	CTTGTTCCAAAGGAAATCCTGGCAGAGGACAAAGTCTTTGTGATATGCAAAATGCTGGT	121
QY	381	AepSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal	400
Db	1219	GATTCGAAAGCTCCGACGAATATCACTGACGTGGAAGATGTTTATATGGGAAGTTTGTTGTG	127
QY	401	AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu	419
Db	1279	GCTACATTCTACCCGCTGAGTGGCTCTACAAAGGGAGATATCTGTATCAAACTTGCTT	1335
RESULT 5			
AA166828			
ID	AA166828	standard; cDNA; 2622 BP.	
XX	AA166828;		
XX	07-JAN-2002	(first entry)	
DT			
XX			
DE		Human protein kinase polypeptide 13887 encoding cDNA.	
KW		Protein kinase; 3714, 16742, 23546, 13887; cancer; bone disorder; human;	
KW		cytotoxic; antinflammatory; immunosuppressive; cardiact; hepatotropic;	
KW		gene therapy; ss.	
XX			
OS		Homo sapiens.	
FH			
FT	Key	Location/Qualifiers	
FT	5'UTR	1..266	
FT		/*tag= a	
FT	CDS	267..1529	
FT		/*tag= b	
FT		/product= "protein kinase 13887"	
FT		/note= "coding sequence specifically claimed"	
FT	3'UTR	1530..2622	
FT		/*tag= c	
XX			
PN		WO200173050-A2.	
XX			
PD		04-OCT-2001.	
XX			
PF		23-MAR-2001; 2001WO-US009483.	
XX			
PR		24-MAR-2000; 2000US-0191846P.	
XX			
PA		(MILL-) MILLENNIUM PHARM INC.	
XX			
PI		Meyers R;	
XX			
DR		WPI: 2001-611632/70.	
XX		P-PSDB; AAG65767.	
PT			
PT		New human protein kinase polypeptides, 3714, 16742, 23546 and 13887,	
PT		useful in diagnosis of cancer or cellular proliferation or	
PT		differentiation disorders and to screen for polypeptide modulators useful	
PT		to treat such conditions.	
XX			
PS		Claim 1; Fig 10A-C; 169pp; English.	
XX			
CC		The invention provides novel human protein kinase polypeptides, 3714,	
CC		16742, 23546 and 13887 and nucleic acid molecules encoding them. The	
CC		protein kinase polypeptides can be expressed by standard recombinant	
CC		methodology. 3714, 16742, 23546 or 13887 nucleic acids and polypeptides	
CC		are useful for diagnostic and screening methods to identify subjects (at	
CC		risk of) having cancer or cellular proliferation and/or differentiation	
CC		disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and	
CC		modulators are useful for the treatment of cancer, particularly colon	
CC		cancer or cellular proliferation and/or differentiation disorders. Other	

CC	disorders associated with 3714, 16742, 23546 or 13887 expression or activity that can be treated include bone related disorders, inflammatory disorders, autoimmune diseases, cardiovascular disorders and liver diseases.						
CC	The present sequence represents a human protein kinase polypeptide 13887 encoding cDNA						
CC							
XX							
Seq	Sequence 2622 BP; 652 A; 567 C; 627 G; 776 T; 0 U; 0 Other;						
<hr/>							
Alignment Scores:							
Pred. No.:	5.11e-206	Length:	2622				
Score:	2195.00	Matches:	417				
Percent Similarity:	99.76%	Conservative:	1				
Best Local Similarity:	99.52%	Mismatches:	1				
Query Match:	99.50%	Indels:	0				
Dbs:	4	Gaps:	0				
<hr/>							
US-10-798-532-4 (1-419) x AA166828 (1-2622)							
OY	1	MetaIagiSerGIyCySaIatRGlVAlAGluProPoaArgPhenLeugIuaIaphGly	20				
Db	267	ATGGCGGGATCCGGCTCGCCTGGGGCGCGAAGCCGCCGTTCCTGGAGGCCTTGGG	326				
OY	21	ArgLeutRpgIuaIgiInserArgLeugLySerGIySerSeraIaservaiYraGvaI	40				
Db	327	CGGCTGTGGCAGGTACAGAAGCCGTCTGGGRAGCGGCTCTCCGCCCTCGTGTAACGGGTT	386				
OY	41	ArgCySGvGIyaAnPrOGlySeRPProPGIyVAIAleuArGiInphLeuProPGIy	60				
Db	387	CGCTGCTGGGCAACCCTGGCTCCGCCCGGCCCTCAAGCATGTTCTTGGCCGCAGA	446				
OY	61	ThrThrgIyaIaIaaIservAIaGlVTrGIyPheaArgLySGIuaArgAIaIaleu	80				
Db	447	ACCACGGGGCTGGGCTCTGGCCGCGAGTAAGTTCCGCAAGAGAGGGCGGCTG	506				
OY	81	GIuGIuLeuGIyIHiaArgAsnIleValThrLeutyRGIyAlPheThrIleHisPhe	100				
Db	507	GAAcAGTTGcAGGGTcAcAGAAcATGcGTGcATTGTrAGGAGcGTTTAcATcCACATT	566				
OY	101	SerProAnaIProSerARgCyLeuLeuLeuGIuLeuLeuAPVaIServAISerGIu	120				
Db	567	TCTCCAAATGGCCATACCGCTGTCTGTGTGTAACCTCGAATGTCAGTGTTGGAA	626				
OY	121	LeuLeuLeuTyRSerSerHieGLIGLyCySerMetTRPmetIleGINHisCYaIARG	140				
Db	627	TTCCTTTATTTCCAgtCACcAGGGTTTCCAtGTGGATGARAcAGCATGTGGCCGA	686				
OY	141	ASPValLeuGIuaIaleuAIaphLeuHiHisGLIGLyTYrValHIaIaaRPLeuLyS	160				
Db	687	GATGTTTTGGAGGCCCTTGCTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAA	746				
OY	161	ProArgAnaIIeLeuTPSRserAlaGIuAbngLIuCyShelysLeuIIeaRPPhGIyleu	180				
Db	747	CCACGTACAATATGTGGAGTGGAGAGAAGATGAAGTTTAAACCTATGACTTTGGACCTT	806				
OY	181	SerPhelysGIuGIYAengIaASPValIystrITileGINtraSPGIyTYrARGAIPro	200				
Db	807	AGCTTCANAAGAGGCATCAGAGTAGTAAGTATTTCAAGACAGCGGGTATCGGGCTCCA	866				
OY	201	GIuIaGIuleuGIaSnCYaLeuAIaGIaIaGIyleuGIuseRaRPthrgIuCyStrHr	220				
Db	867	GAAcAGAAHTTGCAAAATTTGCTTGGCCCAAGGCTGGCTGGAGAATGATACAGATGTACC	926				
OY	221	SerAlaValaAPLeuTPSRserLeuGIyIleIIeuleuGIuWetPhieSerGIymetLys	240				
Db	927	TCAGCTGTGTATCTGTGGAGCTTGAAGTCAATTATTCTGGAATAAGTTCTCGAGATCAA	986				
OY	241	LeuIySHIstrIVaIArgSerGIuITrpIyAlaAnseRSeraIaIeIASPHis	260				
Db	987	CTGAACACTACAGTCOAgaTCTCAAGAAITGGAAAGCAACAGTTTGCTATTATTGATCAC	1046				
OY	261	IlePheaIseryIaIaValaIaASnaIaIaIleProAlaTYRHIEuArgASPLEu	280				
Db	1047	ATATTTCGCAATAAAGcAGTGTGAATGGCCGCAATTCAGCTATCACTTAAGAGACTT	1106				

```
QY 281 ILeYSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
DB 1107 ATCAAAAGCATGCTTCATGATGATCCAGAGAGAGAAATCTCTGTAATGCGATTGTGC 1166
QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
DB 1167 AGCCCATTTCTTATGACATTCCTTTTGGCCCTCATATTGAAGATCGGTGCATGCTCCCACT 1226
QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrIleuGluAsnGluGluTyr 340
DB 1227 CCAAGTCTAAGATCGCTGATGATGCTGATGATGATGATGATGATGATGATGATGATGAT 1286
QY 341 GluAspValValGluAspValValGluGluGluGluGluGluGluGluGluGluGluGlu 360
DB 1287 GAAGATGTTGTATGAAGATGTTGAAGAGAGATGTTCAAAAATATGACAGCGGTATCTCTA 1346
QY 361 LeuValProlysgluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
DB 1347 CTGTTCCAAAGGAAATCTCGGACAGAGACAAGTCTTTGTGATGATGCAAAATGCTGGT 1406
QY 381 AspSerIleValAlaGlnIleuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
DB 1407 GATTCCAAAGCTGCGAGAAATTACTGACTGAGAGAGATGTTGATGGAGAGTTGTGTG 1466
QY 401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrIleuTyrGlnThrLeuLeu 419
DB 1467 GCTACATTTCTACCCGCTGAGTGCCTACAGAGGGGATATCTGATCAAACTTGCTT 1523

RESULT 6
AAH72714
ID AAH72714 standard; cDNA; 4065 BP.
XX
XX AAH72714;
XX
XX 19-SEP-2001 (first entry)
XX
XX Human cervical cancer marker nucleic acid 3988.
XX
XX Cervical cancer; cytostatic; pre-malignant condition; gene therapy; 88.
XX
XX Homo sapiens.
XX
XX MO200142467-A2.
XX
XX 14-JUN-2001.
XX
XX 08-DEC-2000; 2000MO-US033312.
XX
XX 08-DEC-1999; 99US-0169681P.
XX
XX 21-DEC-1999; 99US-0171350P.
XX
XX 14-MAR-2000; 2000US-0189315P.
XX
XX 12-MAY-2000; 2000US-0203791P.
XX
XX 09-JUN-2000; 2000US-0210600P.
XX
XX 21-JUL-2000; 2000US-0220114P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Deeds J, Berger A, Zhao X;
PI
XX WPI; 2001-375006/39.
XX
XX New isolated nucleic acid for diagnosing and treating cervical cancer and
PT for assessing and detecting compounds for treating the cancer.
XX
XX Claim 1; Page 810-811; 1051p; English.
XX
XX The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
```

```
CC Inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy
XX
SQ Sequence 4065 BP; 1068 A; 822 C; 905 G; 1265 T; 0 U; 5 Other;

Alignment Scores:
Pred. No.: 9,696-206 Length: 4065
Score: 2195.00 Matches: 417
Percent Similarity: 99.76% Conservative: 1
Best Local Similarity: 99.52% Mismatches: 1
Query Match: 99.50% Indels: 0
DB: 4 Gaps: 0

US-10-798-532-4 (1-419) x AAH72714 (1-4065)
QY 1 MetAlaGlySerGlyCysAlaAlaTrrGlyAlaGluProProArgPheLeuGluAlaPheGly 20
DB 268 ATGGCGGGATCCGGCTGCGCCCTGGGGCGCGAGCCGCCGTTTCTGGAGGCTTCGGG 327
QY 21 ArgLeuTrrGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
DB 328 CCGCTTGGCAGATACAGAGCCCTGCGGTACCGGCTCTCCGCTCGGTATCGGTT 387
QY 41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProGly 60
DB 388 CGCTGCTGGGCAACCCCTGCTGCCGCCGCCGCCCTCAACAGATTTCTGCCCGCAGGA 447
QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
DB 448 ACCACCGGGGCTGGCGCCCTCTGCCCGCGATGATGTTCCGAAAGAGGGCGGCTG 507
QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
DB 508 GAAACGTTCAAGGTCACAGAAACATGTAAGCTTTGTATGAGCTTTTCAATCCACTT 567
QY 101 SerProAsnValProSerArgCysLeuLeuGluGluLeuAspValSerValSerGlu 120
DB 568 TTCCAAAATGTCATCAACGCTGTGTTGCTTGAAGCTCTGAGATGCAGTTCGGAA 627
QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrrMetIleGlnHisCysAlaArg 140
DB 628 TTGCTCTTATATTCACAGTCACAGAGGTGTGTTCCATGTGATATCAGATATGTC 687
QY 141 AspValLeuGlnAlaLeuAlaPheLeuHisGlnGluTyrValHisAlaAspLeuLys 160
DB 688 GATGTTTGGAGCCCTTGCTTTCTTCATCAAGAGGCTATGTCATCGGACCTCAAA 747
QY 161 ProArgAsnIleLeuTrrPserAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
DB 748 CCAAGTAAATATTTGGAGTGCAGAGAAATGAATGTTTAAATCATTCATCTTGGACTT 807
QY 181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
DB 808 ACCTTCAAAAGAGGCAATCAGAGATGAAGTATATTCAGACAGAGGGGTATGGGCTCCA 867
QY 201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspPheGlyCysThr 220
DB 868 GAAGCAGAAATTCAAATATGCTTGAGCCAGGCTGGCTGCAGAGATGATCAAGAAATGACC 927
QY 221 SerAlaValAspLeuTrrPserLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
DB 928 TCAGCTGTGATCTGTGGAGCTTGAAGATCATTTTAACTGGAATGTCTCAGAGATGAAA 987
QY 241 LeuLysHisThrValArgSerGlnGluTrrLysValAsnSerSerAlaIleLeuAspHis 260
DB 988 CTGAAGACATACAGTCAAGATCTCAGAGATGAGGCAACAGTTCTGCTATTTATATCAC 1047
QY 261 IlePheAlaSerLysAlaValAlaAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
DB 1048 ATATTTTCCAGTAAAGCATGCTGTAATGCCGAATTTCCAGCTTATCACCTTAAGAACCTT 1107
QY 281 ILeYSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
```

Db	1108	ATCAAAAGCATGCTTCAGATGATTCCAAGCAGAAAGATTCCTGCTGAAATGGCATTTGGC	1167
Qy	301	SeRPProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr	320
Db	1168	AGCCCATCTTTAGCATCTCTTTGGCCCTCATATTGAAGATCTGGTCATCTCCCACT	1227
Qy	321	ProValIleuArgLeuLeuAsnValIleuAspAspPheTyrLeuGlyAsnGluGluTyr	340
Db	1228	CCAGTGTCTMAACCTGCTGAATGTGCTGATGATGATTTCTTGGAAATGAAGAGAAAT	1287
Qy	341	GluAspValValIleGluAspValLysGluGluLysCysGluLysTyrGlyProValValSerLeu	360
Db	1288	GAAAGTGTGTGAAGATGTAAAGAGAGAGTCTCAAAATATGGACCAGTGTATCTCTA	1347
Qy	361	LeuValProLysGluAsnProGlyArgGlyGluValPheValGluTyrAlaAsnAlaGly	380
Db	1348	CTTGTTCCAAAGGAAAACTCTGGCGCAGAGACAAGCTTTGTTAGTAGCAAAAGCTGGT	1407
Qy	381	AspSerLysValAlaIleGluLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal	400
Db	1408	GATTCCAAAGCTGGCGCAGAAATTACTGACTGGAAGAGATTGGATGGGAACTTTGTTGTG	1467
Qy	401	AlaThrPheTyrProLeuSerAlaTyrTyrArgGlyTyrLeuTyrGlnThrLeuLeu	419
Db	1468	GCTACATTCTACCCGCTGAGTGCCTACAAAGAGGGGATTCGTATCAAACTTGCTT	1524

the patient an antisense oligonucleotide complementary to a marker of the invention. The markers are useful for assessing if a patient is afflicted with ovarian cancer, which involves comparing the level of expression of a marker in a patient sample and a normal level of expression of the marker in a control non-ovarian cancer sample. A difference between the expression levels indicates ovarian cancer. The level of expression of a marker corresponds to a secreted protein or to a transcribed polynucleotide or its portion. The level of expression of the marker is assessed by detecting the presence in the sample, a protein or protein fragment corresponding to the marker. The presence of protein or protein fragment is detected using an antibody that specifically binds with the protein or protein fragment. Alternatively, the level of expression of the marker is assessed by detecting the presence of a transcribed polynucleotide which anneals with the marker or anneals with a portion of the polynucleotide comprising the marker, under stringent conditions. The marker is also used for monitoring the progression of ovarian cancer in a patient which involves detecting expression of the marker in a patient sample at a first point in time, repeating the method at a subsequent time and comparing the level of expression. The method is carried out using an ovarian tissue sample. A composition comprising a marker, polypeptide or antibody of the invention is used to treat ovarian cancer. This sequence represents a human ovarian cancer DNA marker of the invention.

Sequence 4065 BP; 1068 A; 822 C; 905 G; 1265 T; 0 U; 5 Other;

Alignment Scores:

pred. No.:	9,696-206	Length:	4055
Score:	2195.00	Matches:	417
Percent Similarity:	99.76%	Conservative:	1
Best Local Similarity:	99.52%	Mismatches:	1
Query Match:	99.50%	Indels:	0
DB:	5	Gaps:	0

US-10-798-532-4 (1-419) x ADL45238 (1-4065)

QY	1	MetAlaGlySerGlyCyValaIrrProGlyValaGluProPrcArgPheLeuGluAlaPheGly	20
Db	268	ATGGCGGGATCCCGGCTGGCCCTGGGGGCGCGAGCGCGCTTTTCTGGAGCGCTTGGG	327
QY	21	ArgLeuTrpGluValGlnSerArgLeuGlySerGlySerSerAlaSerValYArgVal	40
Db	328	CGGCTGTGGAGGTACAGAGCCGTCCTGGGTACGGCTCTCCGCGCTCGGTATACGGGTT	387
QY	41	ArgCySeGlyYAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly	60
Db	388	CGGTGCTGGGCAACCTGGCTCGGCCCGCGGCTTCAGACAGTTCTTGGCCGACGA	447
QY	61	ThrThrGlyValAlaAlaSerAlaAlaGluTrpGlyPheArgGlyGluArgAlaAlaLeu	80
Db	448	ACCACCGGGCTGGCGGCTCTCGCCCGCAGATAGGTATTCGCCAAGAAGAGCGGCGCTG	507
QY	81	GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTrpGlyValPheThrIleHisPhe	100
Db	508	GAAACAGTTGCAAGGTACAGAAACATCGTACCTTTGTATGAGAGTGTTACATCCACTTT	567
QY	101	SerProAsnValProSerArgCysLeuLeuLeuGlnLeuLeuAspValSerValSerGlu	120
Db	568	TCTCCAAATGTGCATCAGCTGTCTGTGTGCTTGAACCTCGGATGTGCAGTGTTCGAA	627
QY	121	LeuLeuLeuTrpSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg	140
Db	628	TTCTCTTATATTCACAGTACACAGAGGTGTTCATAGTGAAGATACAGCATTTGGCCGA	687
QY	141	AspValLeuGluAlaLeuAlaPheLeuHisHisGlnGlyTrpValHisAlaAspLeuLys	160
Db	688	GAGGTTTGGAGGCCCTTGGCTTTTCTTATCATCTGAGGGCTATGTCCATGCGGACCTGAAA	747
QY	161	ProArgAsnIleLeuTrpSerAlaGluAsnGlnCysPheLysLeuIleAspPheGlyLeu	180
Db	748	CCACGTACATATTGTGGAGTCCAGAGATGAATTTTAAATCATTTGACTTTGGACTT	807
QY	181	SerPheLysGlnGlyYAsnGlnAspValIleTrpIleGlnIrrAspGlyTrpArgAlaPro	200


```

Db      808 AGCTTCAAGAGGAGCATGAGATGTAAGTATATTCAAGACGAGTATGCGGTCCA
Qy      201 GUAAGAGUleuGlnAAsnCySleuAlaGlnAlaGlyLeuGlnSerPheThrGlyCysThr
Db      868 GAGCGAAGATTGGGAAATTCCTTGGCCCGAGCTGGCTCGACAGATATCAGAAATGACC
Qy      221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuGlnLysPheSerGlyMetLys
Db      928 TCAAGCTTTGATCTGTGAGAGCTTGAAGATCATTTTCTGGAAATGTTCTCAGAAATGAAA
Qy      241 LeuLysHisThrValArgSerGlnIuTrpLysAlaAsnSerSerAlaIleIleAspHis
Db      988 CTGAAGCATATACATGATCATCTGAGAAATGAGAAAGCAAGTTCTGCTATTATGATCAAC
Qy      261 IlePheAlaSerLysAlaValAlaAsnAlaAlaIleProAlaTyrHisSleuAlaPheLeu
Db      1048 AATATTGGCCAGTAAACAGAGTGGTGAATGCCGCAATTCAGCCCTTATCAAGACACCTT
Qy      281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGlnMetAlaLeuCys
Db      1108 ATCAAAAGCATGCTTATGATGATCCAGACGAAAGAAATTCCTGCTAAATGGCATTTGTC
Qy      301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr
Db      1168 AGCCCATTTCTTATGATCTCTTTGGCCCTCATATGGAATGCTGATGCTTCCCACT
Qy      321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluTyr
Db      1228 CCAGTCTTAAGCATCTGCAATGCTGATGATGATTAATCTTGAGATGAAGAGCAATAT
Qy      341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValAlaSerLeu
Db      1288 GAAAGATGTTGAGAAATGTAAGAGAGGTGCAAAATATGACACAGGATCTCTA
Qy      361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly
Db      1348 CTTGTTCCAAAGGAAATCTCTGCGACAGGACAAAGTCTTTGTTGATGAAATGCTGGT
Qy      381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyLysMetPheAspGlyLysPheValVal
Db      1408 GATTCCAAAGCTGCGCAGAAATTAATCACTGGAAGAGATGTTGATGGAAGTTTGTGTG
Qy      401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrIleuTyrGlnThrLeuLeu
Db      1468 GCTACATTTCACTCCGCTGAGTGCCTCAAGAGGGGATATCTGTATCAAACTTGTCTT
RESULT 8
AAS06734
ID      AAS06734 standard; cDNA; 1260 BP.
XX      AAS06734;
AC      AAS06734;
DT      12-SEP-2001 (first entry)
XX      XX
DE      Polynucleotide sequence encoding human protein kinase #34.
XX      XX
KW      Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW      metabolic disorder; immune related disease; neurological disorder;
KW      neurodegenerative disorder; inflammatory disorder; infectious disease;
KW      reproductive disorder; gene therapy; ss.
XX      XX
OS      Homo sapiens.
XX      XX
PN      WO200138503-A2.
XX      XX
PD      31-MAY-2001.
XX      XX
PF      22-NOV-2000; 2000MO-US032085.
XX      XX
PR      24-NOV-1999; 99US-0167482P.
XX      XX
PA      (SUGEN-) SUGEN INC.

```

```

XX      XX
PI      Plowman GD, Whyte D, Manning G, Sudarshanam S, Martinez R;
PI      Flanagan P, Clary D;
XX      XX
DR      WPI; 2001-343950/36.
XX      XX
PT      Nucleic acids encoding human kinase polypeptides, useful for preventing
PT      diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT      neuronal-associated diseases, and microbial infections.
XX      XX
PS      Example 1; Fig 1; 433p; English.
XX      XX
CC      AAS06701-AAS06757 encode for novel human protein kinases #1-57. The novel
CC      protein kinases have been identified as members of the tyrosine or
CC      serine/threonine kinase (PTK and STK) families. The polynucleotides
CC      encoding protein kinases and the polypeptides may be used in the
CC      prevention, diagnosis and treatment of diseases associated with
CC      inappropriate kinase expression. For example, they may be used to treat
CC      cancers (especially cancers of haematopoietic origin), cardiovascular
CC      disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC      immune related diseases (e.g. rheumatoid arthritis), neurological
CC      disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC      Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC      disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC      Additionally, polynucleotides encoding protein kinases may be used for
CC      gene therapy and as DNA probes in diagnostic assays. The protein kinase
CC      polypeptides may be used as antigens in the production of antibodies
CC      against the protein kinases and in assays to identify modulators of
CC      protein kinase expression and activity
XX      XX
SQ      Sequence 1260 BP; 311 A; 279 C; 335 G; 335 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      8,59e-206      Length:      1260
Score:          2188.00      Matches:      416
Percent Similarity: 99.52%      Conservative: 1
Best Local Similarity: 99.28%      Mismatches: 2
Query Match:    4          Gaps:      0
DB:             4          Indels: 0

US-10-798-532-4 (1-419) x AAS06734 (1-1260)
Qy      1 MetAlaGlySerGlyCysAlaIleTrpGlyAlaGluProPheLeuGlnAlaPheGly 20
Db      1 ATGGCGGATTCGCGCTGCGCTGGGAGCGCGCGCGGCTTCTGAGGCTTCGGG 60
Qy      21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
Db      61 CGGCTGTGGCAGATGACAGACCGCTCTGGGTAGCGGCTCTCCGCTCGGTATGGGTT 120
Qy      41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProGly 60
Db      121 CGCTGCTGGCGCAACCTGGCTGCGCCCGCGCGCTCAAGCAATGTTCTTCCGCGCAGGA 180
Qy      61 ThrTrpGlyAlaAlaAlaSerAlaIleGlnTyrGlyPheAspTrpGluArgAlaLeu 80
Db      181 ACCACCGGGGCTGGCGCTCTGCGCCGAGTATGATTTCCGAAAGAGGGCGGCGCTG 240
Qy      81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
Db      241 GAAAGCTTCAAGGATCAAGAAACATCGTGAATTTGATGAGATGTTTCAATCCACTT 300
Qy      101 SerProAsnValProSerArgCysLeuLeuGlnLeuLeuAspValSerValSerGlu 120
Db      301 TCTCCAAATGTCATCAAGCTGTCTGTGATCTGGAATCTCGAGATGATGTTTCGGA 360
Qy      121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysValArg 140
Db      361 TTGCTTTATATTTCAAGTACCAAGGATGTTTCAATGATGATATCAGCATTTGCTCCGA 420
Qy      141 AspValLeuGluAlaLeuAlaPheLeuHisGluGlyTyrValHisAlaAspLeuLys 160

```



```
Db 169 ATGGCGGATCCGGCTGGCGCTGGGGCGGAGACCGCGGCTTTTCGAGGGCTTCGGG 228
Qy 21 ArgLeuThrGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
Db 229 CGGCTGTGGCAGGTACAGAGCCGTGTGGGTAGGGCTCCCTCCCGCGGTGTATCGGGTT 288
Qy 41 ArgCysCysGlyAsnProGlySerProGlyAlaLeuArgGlnPheLeuProGly 60
Db 289 CGGTGTGGGCAACCTGGCTGGCCCCCGGGCCCTCAAGCATTTCTTGGCCGACGA 348
Qy 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgGlyGluArgAlaLeu 80
Db 349 ACCACCGGGGCTGGCGCTTCCTGGCGGATGTTTCGCAAGAGAGGGCGGCGCTG 408
Qy 81 GluGlnLeuGlnGlyIleArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
Db 409 GAACAGTTCCAGGGGTACAGAAACATCTGTACTTGTATGAGCTGTTTCAATCCACTTT 468
Qy 101 SerProAsnValProSerArgCysLeuLeuLeuGlnLeuAspValSerValSerGlu 120
Db 469 TCTCCAAATGGCCATCAGCTGTCTGTGCTTGAACCTCTGGATGTCAAGTGTTCGAA 528
Qy 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
Db 529 TTGCTCTTATATCCAGTACAGAGGTTGTTCATGTGATGATACAGCATTTGTGCCCA 588
Qy 141 AspValLeuGlnAlaLeuAlaPheLeuHisGlnGlyTyrValHisAlaAspLeuLys 160
Db 589 GATGTTTGGAGGCCCTTGTCTTCTTCTTCATCATGAGGGCTATCTCATGCGGACCTCAAA 648
Qy 161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLeuLeuLeuAspPheGlyLeu 180
Db 649 CCACGTAACATATTTGGAGTGCAGAGATGATGATTTTAACTCATTTGACTTTGACCTT 708
Qy 181 SerPheLeuGlnGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
Db 709 AGCTTCAAAGAGGAGCATCAGATGATGATTAATTCAGACAGACGGGTATCGGGCTCA 768
Qy 201 GluAlaGlnLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGlyCysThr 220
Db 769 GAGCGAATTTGCMAATTCCTTGGCCAGGGCTGCAGAGTATCAGATATGATACC 828
Qy 221 SerAlaValAspLeuTrpSerLeuGlyIleLeuLeuGlnLeuPheSerGlyMetLys 240
Db 829 TCAGCTGTGATGTGAGGCTTGAAGATATTTCTGAAATGTTCTCAGAAAGAA 888
Qy 241 LeuLysHisThrValArgSerGlnGluTyrLysAlaAsnSerSerAlaIleIleAspHis 260
Db 889 CTGAACATACAGTCAAGATCTCAGAAATGAGAGGCAACAGTTCTCTATTAATGATCAC 948
Qy 261 IlePheAlaSerLysAlaValAlaValAlaIleProAlaTyrHisIleuArgAspLeu 280
Db 949 ATATTTTCCAGTAAACAGTGGTGAATGCCGCAATTCAGCTCATACCTTAAGACCTTT 1008
Qy 281 IleLysSerMetLeuHisAspAspProSerArgTyrIleProAlaGlnMetAlaLeuCys 300
Db 1009 ATCAAAAGCATGCTTATGATGATCCAGAGAAAGAAATTCCTGCTGAATAGGCAATGTGC 1068
Qy 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
Db 1069 AGCCCATTTCTTATGATCTCTTTGGCCCTCATATGAAGATGTGGTCAATGCTTCCCACT 1128
Qy 321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlnAsnGluGlnGlyTyr 340
Db 1129 CCAGTCTTAAGACTGTGATGTCTGATGATGATGATTAATCTTGAAGATGAAGGAGATAT 1188
Qy 341 GluAspValValGluAspValLysGlnGluCysGlnLysTyrGlyProValIleSerLeu 360
Db 1189 GAAGAGTGTGTAAAGATGTAAAGAGAGATGTCAAAATATGACACAGTGGTATCTCA 1248
Qy 361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
```

```
Db 1249 CTTGTTCMAAGGAAATCTTGCGAGAGACAGTCTTTGTGATGATCAAAATCTGGT 1308
Qy 381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
Db 1309 GATTCMAAGCTGCCGAGAAATTAATCTGAGAGAGATGTTGATGGAGAGTGTGTG 1368
Qy 401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
Db 1369 GCTACATTTACCCCTGATGCTCTACAGAGGGATATCTGTATCAAACTTGCTT 1425

RESULT 10
ADBS2857
ID ADBS2857 standard; DNA; 3244 BP.
XX
AC ADBS2857;
XX
DE 04-DEC-2003 (first entry)
XX
KW Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3399.
KW toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW toxicity marker; toxicity progression; drug screening;
KW primary rat hepatocyte toxicity modelling; gene; db.
XX
XX Rattus norvegicus.
XX
PN W02003065993-A2.
XX
PD 14-AUG-2003.
XX
PF 04-FEB-2003; 2003W0-US003482.
XX
PR 04-FEB-2002; 2002US-0353171P.
PR 13-MAR-2002; 2002US-0363534P.
PR 08-APR-2002; 2002US-0370248P.
PR 10-APR-2002; 2002US-0371134P.
PR 10-APR-2002; 2002US-0371135P.
PR 10-APR-2002; 2002US-0371150P.
PR 11-APR-2002; 2002US-0371413P.
PR 19-APR-2002; 2002US-0373601P.
PR 19-APR-2002; 2002US-0373602P.
PR 22-APR-2002; 2002US-0374139P.
PR 08-MAY-2002; 2002US-0378370P.
PR 09-MAY-2002; 2002US-0378652P.
PR 09-MAY-2002; 2002US-0378653P.
PR 09-MAY-2002; 2002US-0378655P.
PR 09-JUL-2002; 2002US-0394230P.
PR 09-JUL-2002; 2002US-0394253P.
PR 04-SEP-2002; 2002US-0407688P.
PR 28-JAN-2003; 2003US-0442900P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
PI Elashoff M;
XX
DR WPI; 2003-731472/69.
XX
XX
PT Determining if a compound induces a toxic effect on a tissue or cell, for
PT identifying hepatotoxic compounds, comprises comparing a gene expression
PT profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT mean values.
XX
PS Claim 44; SEQ ID NO 3399; 874bp; English.
XX
XX The present invention describes a method for determining whether a
XX compound induces a toxic effect on a tissue or cell. The method comprises
XX preparing a gene expression profile of a tissue or cell sample exposed to
XX the compound, and comparing the gene expression profile to a database
XX comprising data or information on the Tox mean and non-Tox mean value.
XX The method is useful for predicting or identifying at least one toxic
XX effect, particularly hepatotoxicity, of a test or unknown compound. The
XX genes listed in the specification are useful as diagnostic or toxicity
```

CC markers for the prediction or identification of the physiological state
 CC of tissue or cell sample that has been exposed to a compound, or to
 CC identify or predict the toxic effects of a compound or an agent. The
 CC may also be used as markers for monitoring toxicity progression or for
 CC drug screening. The present sequence represents a primary rat hepatocyte
 CC toxicity modelling related gene sequence from the present invention.

XX Sequence 3244 BP; 800 A; 710 C; 819 G; 915 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1, 07e-204	Length:	3244
Score:	2183.00	Matches:	414
Percent Similarity:	99.52%	Conservative:	3
Best Local Similarity:	98.81%	Mismatches:	2
Query Match:	98.96%	Indels:	0
DB:	10	Gaps:	0

US-10-798-532-4 (1-419) x ADBS2857 (1-3244)

```

QY 1 MetAlaGlySerGlyCybaLaATPGLyAlaGluProPhaRphLeuGluAlaPheGly 20
DB 162 ATGGCCGGGTCGGGTCGGGTCGGGTCGGGTCGGGTCGGGTCGGGTCGGGTCGGG 221
QY 21 ArgLeuTrpGlnAlaGlnSerArgLeuGlySerSerAlaSerValTYrArgVal 40
DB 222 CGGCTGTGGCGAGTACAGAGCCCGCTAAGGCGGGCTCCCTCGGCTCGGTGACCGGGTG 281
QY 41 ArgCybaGlyValAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
DB 282 CGCTGCTGGCGGACTCGAGGCTGCCCCCGGCGCCCTCAAGAGATTCCTGCTCCGGGA 341
QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTYrGlyPheArgGlyGluArgAlaLeu 80
DB 342 ACCACCGGGGCTCCGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGG 401
QY 81 GluGlnLeuGlnGlnHisArgGlnLeuValThrLeuTYrGlyAlaPheThrLeuHisPhe 100
DB 402 GAGCAGATTGGAGGATCAGAGAAATCGACTTTTATACGAGATCTTTACCATACACTTC 461
QY 101 SerProAsnValProSerArgCybaLeuLeuGluGlnLeuAspValSerValSerGlu 120
DB 462 TCTCCAAATGTGCATACAGCTCTGCTTGTGAACTCCTGAGATGTCAGCGTTTGGAA 521
QY 121 LeuLeuLeuTYrSerSerHisGlnGlyCybaSerMetTrpMetLeuGlnHisCybaLaArg 140
DB 522 TTGCTGTGATATTCAGTCCAGGCTGCTCATGATGATGATCCAGCATGTGTGCCAG 581
QY 141 AspValLeuGlnAlaLeuAlaPheLeuHisGlnGlyTYrTYrAlaHisAlaAspLeu 160
DB 582 GATGTCTGGAGGCGCTTGTCTTTTCTTCAATGAGGCTATGTCCATGCAAGCTCAAG 641
QY 161 ProArgAsnLeuLeuTrpSerAlaGluAsnGluCybaPheLeuLeuAspPheGlyLeu 180
DB 642 CCACGAAATCTCTGTGAGTGGCGGAGAACGATGCTTTAAGCTTTATGACTTTGAACTC 701
QY 181 SerPheLeuGluGlyAsnGlnAspValLYeTYrTLeuGlnThrAspGlyTYrArgAlaPro 200
DB 702 AGCTTCAAGAGGCAATCAGACGTGAAGTATATTCAGACAGCGGATATCAGACTCTCT 761
QY 201 GluAlaGlnLeuGlnAsnCybaLeuAlaGlnAlaGlyLeuGlnSerSerPheThrGlu 220
DB 762 GAAGCGAAGACTGAGAAATGCTTGGCCGAGCGGCTGAGAGTATCAGAGTGTACC 821
QY 221 SerAlaValAspLeuTrpSerLeuGlyLeuLeuLeuGlnLuecPheSerGlyLuecLys 240
DB 822 TCGAGCTGTGATCTTCGAGCCTTGAATCATTTTACTGGAATTTCTTCAGAAAGAAA 881
QY 241 LeuLYeHisLeuThrValArgSerGlnGluTYrLYeAlaAsnSerSerAlaLeuLeuAspHis 260
DB 882 CTGAACATATCAGATCTCAGAGTGAAGGCAAAAGTTCTGTATATATGATCAT 941
QY 261 IlePheAlaSerLYeValValAlaAsnAlaAlaIleProAlaLYrHisLeuAlaAspLeu 280

```

```

DB 942 ATATTGCCAGTAAAGCAATGCGAATCCGCAATTCACGCTATCAGAGACTT 1001
QY 281 IleLYeSerMetLeuHisAspAspProSerArgAlaIleProAlaGluMetAlaLeuCy 300
DB 1002 ATCAAAAGCATGCTTCATGACGACCAAGAGAAAGATCCCTGCTGAGATGGCTTGAGC 1061
QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
DB 1062 AGCCCATCTTAAAGATTCCTTTTGCCCTCATATTAAGATCTGATGCTTCAACT 1121
QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspArgTYrLeuGlyAsnGluGluTYr 340
DB 1122 CAGTGCTCAGACTCTCTCATATGCTGATGATGACTTCTTGAATAAGAGTAAAT 1181
QY 341 GluAspValAlaGluAspValLYeGluGlyGlnLYeTYrGlyProValAlaSerLeu 360
DB 1182 GAAAGATGTTGTGAAGATGTAAAGAGTGTCAAGAAATATGACCAAGTGGTTCTCTG 1241
QY 361 LeuValProLYeGluAsnProGlyArgGlyGlnValPheValGluTYrAlaAsnAlaGly 380
DB 1242 CTTGTTCMAAGAAATCTTGCGAGAGACAAATCTTGTGTAGTACGCAAAACCTGGA 1301
QY 381 AspSerLYeAlaAlaGlnLYeLeuLeuThrGlyArgMetPheAspGlyLYePheValAla 400
DB 1302 GATTCGAAAGCTGCTCAGAAATGCTGACTGGAGAGATCTTGAAGGAAATTTGTGTG 1361
QY 401 AlaThrPheTYrProLeuSerAlaTYrLYeAspArgLYrLeuTYrGlnThrLeuLeu 419
DB 1362 GCTACATCTTACCCGCTGAGTGTCTTCAAGAGGGATATCTTATCAAACTTGCTT 1418

RESULT 11
ADQ24401
ID ADQ24401 standard; DNA; 3998 BP.
AC ADQ24401;
DT 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 7221.
XX
KW soft tissue sarcoma; cyclostatic; gene therapy; vaccine; screening; human;
de.
OS Homo sapiens.
PN WO2004048938-A2.
XX
PD 10-JUN-2004.
XX
PF 26-NOV-2003; 2003WO-US038193.
XX
PR 26-NOV-2002; 2002US-0429739P.
PA (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Ginsburg WM, Zlotnick A;
XX
PT WPI; 2004-441208/41.
XX
PT Early detection of soft tissue sarcoma comprises determining expression
of a gene in a first soft tissue sample and a normal soft tissue sample
and comparing the gene expression, also useful in treating soft tissue
sarcoma.
XX
PS Example 2; SEQ ID NO 7221; 210pp; English.
XX
CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has

```

CC cytosolic applications and may be useful for detecting soft tissue
 CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
 CC acid sequences may be useful in diagnostic and screening applications.
 CC The current sequence is that of a human soft tissue sarcoma-upregulated
 CC DNA of the invention. The current sequence is not shown within the
 CC specification per se but was submitted in CD format by the inventor.

XX Sequence 3998 BP; 1066 A; 777 C; 885 G; 1255 T; 0 U; 15 Other;

Alignment Scores:

Pred. No.:	1.06e-190	Length:	3998
Score:	2042.50	Matches:	395
Percent Similarity:	94.76%	Conservative:	3
Best Local Similarity:	94.05%	Mismatches:	21
Query Match:	92.59%	Indels:	1
DB:	12	Gaps:	1

US-10-798-532-4 (1-419) x ADQ24401 (1-3998)

QY 1 MetAlaGlySerGlyCysAlaATPGAlaGluProProArpPheLeuGluAlaPheGly 20
 Db 160 ATGGCCGGATCCGGCTGCGCTGGGGCGCGAGCCCGCGTTTCTGAGGCTTCGGG 219
 QY 21 ArgLeuTPGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTYrArgVal 40
 Db 220 CGGCTGTGGAGGTACAGAGCCGCTGGGTAGCGGCTCTCCGCTCGGTGTGGGTT 279
 QY 41 ArgCysCysGlyValanProGlySerProProGlyValAlaLeuArgGlnPheLeuProProGly 60
 Db 280 CGCTGTGGCGCAACCTGTGCTGCCCCCGGCGCCCTCAAGAGATCTTGGCGCAGGA 339
 QY 61 ThrThGlyValAlaAlaSerAlaAlaGluTYrGlyPheArgGlyGluArgAlaAlaLeu 80
 Db 340 ACCACCGGGGCTCGGCTCTGCGCCGAGATGTTTAAAGCATGTTTCTACCAACCG 399
 QY 81 GluGlnLeuGlnGlyHisArgAniLeuValThrLeuTYrGlyValPheThrIleHis 99
 Db 400 AAGGTCTGTGTGACCTTAAACGAGAGTGACTTTGTATGAGATGTTTCAATCCAGGTT 459
 QY 100 PheSerProAnValProSerArgCysLeuLeuLeuGluLeuLeuAnValSerValSer 119
 Db 460 GTTCTCCAAAGTGTCATCAACCTGTGCTGCTTGAATCTTGATGATGATGATGATG 519
 QY 120 GluLeuLeuLeuTYrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAla 139
 Db 520 GAATTCCTCTTATATTCAGTCAACGAGGTTGTTCAATGATGATGATGATGATGATG 579
 QY 140 ArgAspValLeuGluAlaLeuAlaPheLeuHisGlnGlyTYrValHisAlaSerLeu 159
 Db 580 CGAGATGTTTGGAGGCCCTTGTCTTTCTTCAATGATGAGGCTATGATGATGATGATG 639
 QY 160 LysPheArgAniLeuLeuTPSerAlaGluAnGlyCysPheLeuValLeuPheGly 179
 Db 640 AAACCAAGTAACTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 699
 QY 180 LeuSerPheLeuGlyGluGlnAnGlyValPheValTYrIleGlnThrArgPheGlyTYrArgAla 199
 Db 700 CTTAGCTTCAAGAGAGCAATCGATGATGATGATGATGATGATGATGATGATGATG 759
 QY 200 ProGluAlaGluLeuGlnAnGlyCysLeuAlaGlnAlaGlyLeuGlnSerArpThrgLys 219
 Db 760 CCAGAGACGAATTTGCAAAATTTGCTTGCCCAAGCTGGCTGCAAGTATGATGATGATG 819
 QY 220 ThrSerAlaValAspLeuTPSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMet 239
 Db 820 ACCTCAGCTGTGATCTGTGAGACCTTGAAGATCATTTTATGTAATGTTTCTCAGGAATG 879
 QY 240 LysLeuLysHisThrValArgSerGlnGluTPTrpValAlaAnSerSerAlaIleIleLeu 259
 Db 880 AAATTAACATTCAGTCAATCTCAGAAATGGAAGCAACGTTCTGATATATGAT 939
 QY 260 HisIlePheAlaSerLysAlaValAlaAnAlaAlaIleProAlaTYrHisLeuArgAsp 279

Db 940 CACATATTTCCAGTAAAGCACTGGTGAATGCCGCAATTCACCTTACCTTAAGAGAC 999
 QY 280 LeuIleLysSerMetLeuHisAspArpProSerArgGlyIleProAlaGluMetAlaLeu 299
 Db 1000 CTATATCAAAAGATCTTCATATATATCCAGCAAGAAATTCCTGCTGAATGGCATTG 1059
 QY 300 CysSerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuPro 319
 Db 1060 TGCAGCCATTTTATAGCATTCCTTTGGCCCTCATATTTGAAGATCTGGTATGCTTCCC 1119
 QY 320 ThrProValLeuArgLeuLeuAnValLeuAspAspArgTYrLeuGlyAnGluGlu 339
 Db 1120 ACTCAGAGCTTAAGATCTGTGATGTGTGATGTATTTCTTGAAGATGAAAGGAA 1179
 QY 340 TYrGluAspValValGluAspValLysGluGluCysGlnLysTYrGlyProValAlaSer 359
 Db 1180 TATGAAGATGTGTGAAGATTAAGAGGAGTGTCAAAATATGACCACTGGATCT 1239
 QY 360 LeuLeuValProLysGluAnProGlyValArgGlyGlnValPheValGluTYrAlaAnAla 379
 Db 1240 CTACTTGTCCAAAGAAATCTGCGAGAGACAAGCTTTGTTGATGATGCAATGCT 1299
 QY 380 GlyAspSerLysAlaAlaGlnLysLeuThrGlyArgMetPheAspGlyLysPheVal 399
 Db 1300 GGTGATTTCCAAAGCTGCGCAAAATTAAGTGAAGAGAGTGTGATGAGGAACTTTGTT 1359
 QY 400 ValAlaThrPheTYrProLeuSerAlaTYrLysArgGlyTYrLeuTYrGlnThrLeu 419
 Db 1360 GTGGCTACATTTACCCGCTGAGTCTCAACAAGAGGATATCTGTATCAAACTTGCT 1419

RESULT 12
 AAH70764
 ID AAH70764 standard; cDNA; 529 BP.
 XX
 AC AAH70764;
 DT 19-SEP-2001 (fixet entry)
 XX
 DE Human cervical cancer marker nucleic acid 2038.
 XX
 KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200142467-A2.
 XX
 PD 14-JUN-2001.
 XX
 PF 08-DEC-2000; 2000WO-US033312.
 XX
 PR 08-DEC-1999; 99US-0169681P.
 XX
 PR 21-DEC-1999; 99US-0171350P.
 XX
 PR 14-MAR-2000; 2000US-0189315P.
 XX
 PR 12-MAY-2000; 2000US-0203791P.
 XX
 PR 09-JUN-2000; 2000US-0210600P.
 XX
 PR 21-JUL-2000; 2000US-0220114P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Deeds J, Berger A, Zhao X;
 XX
 DR WPI; 2001-375006/39.
 XX
 PT New isolated nucleic acid for diagnosing and treating cervical cancer and
 XX for assessing and detecting compounds for treating the cancer.
 XX
 PS Claim 1; Page 434; 1051p; English.
 XX
 CC The invention relates to novel genes (AAH68727-AAH73383) associated with
 CC cervical cancer with cytostatic activity. The nucleic acids and encoded
 CC polypeptides are useful to assess if a patient is afflicted with
 CC cervical cancer or has a pre-malignant condition; to monitor the
 CC progression of cervical cancer or a premalignant condition in a patient;

CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy

SQ Sequence 529 BP; 156 A; 102 C; 132 G; 139 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.8e-68	length:	528
Score:	789.00	Matches:	150
Percent Similarity:	95.78%	Conservative:	0
Best Local Similarity:	95.78%	Mismatches:	5
Query Match:	35.77%	Indels:	2
DB:	4	Gaps:	0

US-10-798-532-4 (1-419) X AAH70764 (1-529)

QY	22	ThSeriValaAAspLeuTSPserLeuGlyIleIleLeuLeuGluMetPheSerGluMet	239
Db	26	ACCTCAAGCTTGAATCTGTGGAGCCTAGAATCATTTTACTGGAAATGTTCTCAGGAATG	85
QY	240	LyLeuLeuYpsiIerThValArgSerGlnIuTrpValaAsnSerSerIalIeIeAsp	258
Db	86	AAACTGAACATCACTCACTCAATCTCCAGCAATGGAAGCAAAACAGTTCGTATTAATGAT	145
QY	260	HisIlePheLeuSerSerValaValaValaAsnIalalIeProIalThYHisLeuValAsp	279
Db	146	CACATATTTCCTCAGTAAAGCAGTGGTGAATGCCGCAATTCAGCTTACTCACTTAAGAC	205
QY	280	LeuIleIeYserMetLeuHisaAspProSerArgArgIleProIalGluMetAlaLeu	299
Db	206	CTTATCAAAAGCATGCTTCATGATGATTCACACAGAAAGATCTCCGTGAATGGCATTG	265
QY	300	CysSerProPhePheSerIleProPheIalProHisiIeGluAspLeuValMetLeuPro	319
Db	266	TGAGGCCATCTTTAGCATCTTTTGGCCCTCATTTGAAGATCTGGTATGCTTCCC	328
QY	320	ThrProValLeuArgLeuLeuAsnValLeuAspAspTyrLeuGluYAsnGluGlu	339
Db	326	ACTCCAGTGTGAAGCTGCTGAATGGTGGATGATGATTACTTGAAATGGAAGAGAA	385
QY	340	TyrGluAspValValaGluAspValIleGluGluCysGlnIleYserGlyProValAs	355
Db	386	TATGAAGATGTGTGTAAGATGTAAAGAGGAGAGTCATAAAATATGACCAAGGGAT	445
QY	359	extLeuLeuValProLysGluAsnProGluYArgGluValaPheValGluTrpAlaAsn	379
Db	446	CTCTACTTGCTCAAAAGAAATCTGGCAGAGGAACAGTCTTTGTTAGTATCAAAAG	505
QY	379	IaGlyAspSerLys	383
Db	506	CTGGGATTTCAAG	519

RESULT 13

ID AAA02535 standard; cDNA; 722 BP.

AC AAA02535 ;

DT 19-MAY-2000 (first entry)

DE Human colon cancer cell line polynucleotide sequence SEQ ID NO:2526.

KW Human; colon cancer; tumour; diagnosis; gene expression product; probe;

KW oestrogen receptor-positive breast cancer; therapy;

XX

XX

XX

XX

PF 13-MAY-1999; 99WO-US010602.

PR 14-MAY-1998; 98US-0085426P.

PR 15-MAY-1998; 98US-0085696P.

PR 27-OCT-1998; 98US-0105877P.

PA (CHIR) CHIRON CORP.
 (INVEST) INVEST TWO

11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533

PI Reinhard C, Giese H

PI Leshkowitz D, Kita

DR WPI; 2000-126369/11.

PT Polynucleotide Library

XX **C) 2-й т. База 1020.**

XX 22200010 40 222002716

libraries constructed

CC genes correlated with
CC detecting at least 1

CC sample derived from
CC of the differentials

CC cancerous state of the

CC differentially expressed genes correlated with a cancerous state of a
CC mammalian cell. The polynucleotides can also be used as probes for
CC detecting and mapping related genes. They can be used in diagnosis and
CC prognosis of diseases and disorders (e.g. identification of pre-
CC metastatic or metastatic cancerous states, stages of cancer, or
CC responsiveness of cancer to therapy). This is particularly true for breast
CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
CC negative breast cancer, lung cancer, and colon cancer

Alignment Scores:

Score: _____

Pred. No.:	7,32e-6	Length:	722
Score:	766.56	Matches:	18
Percent Similarity:	84.86%	Conservative:	3
Best Local Similarity:	83.49%	Mismatches:	20
Query Match:	34.75%	Indels:	14
DB:	3	Gaps:	1

US-10-798-532-4 (1-419) X AAA02535 (1-722)

OY	152	GIUGIYTYrValHhAlaIAspLeuYsPfoAlaGanIleLeuTPpSerAlaGluAsnGlu	177
Db	1	GAAGGCTAAAGTCCATCGGNCCTCAAAACACATTAATGTGGAGTGCAGAGAAAGAA	60
OY	172	CysPheIAspLeuIleAspPheGlyLeuSerPheIysGluGlyAsnGlnAspValIysTyr	195
Db	61	TGTTTTAAACTCATTTGACTTTGGACTTAACCTCAAGAAGGCAATCAGATGTAAAGTAT	120
OY	192	IIeGlnThrAspGlyTyrTrpAlaProGluAlaGluLeuGlnAsnCysLeuAlaGlnAla	211
Db	121	ATTCAGACGACCGGATCGGGCTCCAGAAC-GAATTGCAGAAATGCTTGCCACAGCT	177
OY	212	GlyLeuGlnSerAspThrGluCysThrSerAlaValAspLeuTrpSerLeuGlyIleIle	231
Db	180	GGCTGAGAGTATACAGATGTACTCAGCTGTGATCTGTGGAGGCTGAGATCATTT	233
OY	232	LeuLeuGluMetCysSerGlyMetIysLeuIysPheIsthrValArgSerGlnGluTrpIys	251
Db	240	TTACTGGAAATGTTCTCAGAGAAATGAACATACAGTCAGATCTTCAGAGAAATGAA	295

QY 252 AlaAsnSer-SerAlaIleIle-AspHisIlePheAlaSerIysAla--ValValAsnAl 270
 DB 300 GCAACAGTTTCGCTATTATTGATCAGATATTCCAGTAAGCAANTGGTGAATGC 359
 QY 270 aAlaIleProAlaIATyHisIleuArgAspLeuIleIysSerMetLeuHisAspAsp-ProS 290
 DB 360 CCGAATTCAGCCTATCAGCTAANAGACCTTATCAAAAGAGATGCTTCATGATATCCAA 419
 QY 290 eArg-Arg-IleIleProAlaIleAlaIleu-CysSerProPhePheSer-IleProPheA 309
 DB 420 GCGAGGAAGATTTCTNCTGAAATGGCATTTGGTGCANCCATTTTACGCAATTCCTTTTG 479
 QY 309 lAProHisIleIleGluAspLeuValMetLeuProThrProVal-LeuArgLeuAsnVal 328
 DB 480 CCCCTCATATTGAAGATCTGNTCATCTTCCACTCAGTGGCTAAGACTGCTGAATGTG 539
 QY 329 Leu-AspAspAspPyrIleuGluYasnGlu--GluGluTyArgIleuAspValIleGluAspVal 347
 DB 540 CTGGGNTGATGATTAATCTTGAGAAATGAAGAGATTAAGATGTTGTTGAAGATGNT 599
 QY 348 IysGluGluCysGlnIys---TyArgIlePro 356
 DB 600 AAAAGAAAGATGCGCAAAATTTTGACCA 679

RESULT 14
 AAH71966
 ID AAH71966 standard; cDNA, 470 BP.
 AC AAH71966;

19-SEP-2001 (first entry)

Human cervical cancer marker nucleic acid 3240.

Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

Homo sapiens.

WO200142467-A2.

14-JUN-2001.

08-DEC-2000; 2000WO-US033312.

08-DEC-1999; 99US-0169681P.

21-DEC-1999; 99US-0171350P.

14-MAR-2000; 2000US-0189315P.

12-MAY-2000; 2000US-0203791P.

09-JUN-2000; 2000US-0210600P.

21-JUL-2000; 2000US-0220114P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Schlegel R, Deeds J, Berger A, Zhao X;

WPI; 2001-375006/39.

New isolated nucleic acid for diagnosing and treating cervical cancer and for assessing and detecting compounds for treating the cancer.

Claim 1; Page 636; 1051pp; English.

The invention relates to novel genes (AAH68727-AAH73383) associated with cervical cancer with cytostatic activity. The nucleic acids and encoded polypeptides are useful: to assess if a patient is afflicted with cervical cancer or has a pre-malignant condition; to monitor the progression of cervical cancer or a premalignant condition in a patient; and to select and/or assess the efficacy of a compound or therapy for inhibiting cervical cancer in a patient. The nucleic acids may also be useful for gene therapy

Sequence 470 BP; 140 A; 91 C; 114 G; 125 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1,966-62 Length: 470
 Score: 729.00 Matches: 147
 Percent Similarity: 96.71% Conservative: 0
 Best Local Similarity: 96.71% Mismatches: 3
 Query Match: 33.05% Indels: 2
 DB: 4 Gaps: 0

US-10-798-532-4 (1-419) x AAH71966 (1-470)

QY 220 ThrSerAlaValAspLeuTrpSerIleuGlyIleIleLeuGluMetPheSerIleMet 239
 DB 18 ACCTCAGCTGTTGATCTGGAGCTAGGAATCAATTATCTGAATGTCTCAGAAATG 77
 QY 240 IysIleuYHisIleThrValArgSerGlnIuTrpIysAlaAsnSerSerAlaIleIleAsp 259
 DB 78 AAACGAAACATACAGTCATGATCTCAGAAATGAAAGCAACAGTTCGTCTATTGANT 137
 QY 260 HisIlePheAlaSerIysAlaValValAsnAlaAlaIleProAlaTyHisIleuArgAsp 279
 DB 138 CACATATTGGCCAGTAAGCACTGTAATGCCGCAATTCAGCCTATCAGCTAAGAGAC 197
 QY 280 LeuIleIysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaIleu 299
 DB 198 CTTATCAAAAGCATGCTTCATGATGATCCAGCAAGAGATTCGTGAATGCGCAATTG 257
 QY 300 CysSerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuPro 319
 DB 258 TGCAGCCCATTTTAAAGCATTCCTTTGGCCCTCATATTGAAGATCTGCTATGCTTCCC 317
 QY 320 ThrProValIleuArgLeuLeuAsnValIleuAspAspAspPyrIleuGluYasnGluGlu 339
 DB 318 ACTCCAGTCTAAGACTGCTGAATGTGCTGATGATGATTAATCTTGAGAAATGAAGAGAA 377
 QY 340 TyArgIleuAsp-ValIleGluAspValIleGlu-CysGlnIysTyArgIleProValIle 359
 DB 378 TATGAAGATTTGTTGTGAAGATGTAAGAGGAGAGTGTCAAAATATATGACAGCGGATAT 437
 QY 359 eIleuLeuValProIysGluAsnProGluYArg 369
 DB 438 CTCTACTGTGTCAAAGAAATCTCTGGCAGA 469

RESULT 15
 AAH69101

ID AAH69101 standard; cDNA, 461 BP.

AC AAH69101;

19-SEP-2001 (first entry)

Human cervical cancer marker nucleic acid 375.

Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

Homo sapiens.

WO200142467-A2.

14-JUN-2001.

08-DEC-2000; 2000WO-US033312.

08-DEC-1999; 99US-0169681P.

21-DEC-1999; 99US-0171350P.

14-MAR-2000; 2000US-0189315P.

12-MAY-2000; 2000US-0203791P.

09-JUN-2000; 2000US-0210600P.

21-JUL-2000; 2000US-0220114P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Schlegel R, Deeds J, Berger A, Zhao X;

DR WPI; 2001-375006/39.

XX New isolated nucleic acid for diagnosing and treating cervical cancer and
PT for assessing and detecting compounds for treating the cancer.

XX Claim 1; Page 168-169; 1051pp; English.

XX The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy

XX Sequence 461 BP; 134 A; 91 C; 109 G; 127 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2,87e-53	Length:	461
Score:	636.00	Matches:	138
Percent Similarity:	93.33%	Conservative:	2
Best Local Similarity:	92.00%	Mismatches:	5
Query Match:	28.83%	Indels:	6
DB:	4	Gaps:	0

US-10-798-532-4 (1-419) x AAH69101 (1-461)

```
QY 220 ThSerAlaValAspLeuTpsSerLeuGlyIleIleuLeuGluMetPheSerGlyMet 239
    |||||
DB 22 ACCTCAGCTGTATCTGTGAGCCCTAGGAAATCATTTTACTGAAATGTTCTCAGGAATG 81
QY 240 LysLeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAsp 259
    |||||
DB 82 AAACCTGAACATACAGTCAGATCTCAGGAATGGAAGCAACAGTTCTGCTATTATTGAT 141
QY 260 HisIlePheAlaSerLysAlaValAlaAsnAlaIleProAlaTyrHisLeuArgAsp 279
    |||||
DB 142 CACATATTGCGCAGTAAGCAGTGTGAATGCCGCAATTCAGCCTATCCCTAAGAGAC 201
QY 280 LeuIleLysSerMetLeuHisAspAspProSerArgArgIleProIleGluMetAlaLeu 299
    |||||
DB 202 CTTATCAAAAGCATGCTTCATGATGATCCAAAGCAGAAATTCCTGCTGAATGSCATTG 261
QY 300 CysSerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuPro 319
    |||||
DB 262 TGCAGCCCAATCTTTAGCATTCCTTTGCCCTCATAT-GAAGATCTGTCATGCTTCCC 320
QY 320 ThrProValIleuArgLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGlu 339
    |||||
DB 321 ACTCCAGTGTAGACTGCTGGAATGCTGATGATGATTATCTTGAATGAAGAGAA 380
QY 340 TyrGluAsp--ValValGluAspValLysGlu--GluCysGlnLysTyrGlyPro-ValV 358
    |||||
DB 381 TATGAAGATTGTGTGAAGATGTAAAGAGGAGAGGTGTCAAAAATATGACCAAGTGG 440
QY 358 alSerLeuLeuValProLys 364
    |||||
DB 441 TATCTCTACTTTGTCCAA 460
```

Search completed: November 29, 2004, 11:42:36
Job time : 520 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2004, 11:26:47 ; Search time 26 Seconds
(without alignments)
1550.570 Million cell updates/sec

Title: US-10-798-532-4

Perfect score: 2206
Sequence: 1 MAGSGCAGMGPFRFLKAFG.....VATFYPLSAKKRYLYQTLL 419

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1052	47.7	200	2	148615 gene KIS protein -
2	248.5	11.3	948	2	T24445 hypothetical prote
3	240.5	10.9	720	2	A96807 hypothetical prote
4	240.5	10.9	826	2	JC5153 mitogen-activated
5	240	10.9	388	2	S48879 protein kinase SKK
6	238.5	10.8	387	2	T37758 protein kinase skp
7	234.5	10.6	433	2	S37790 probable serine/th
8	233	10.6	369	2	A56492 protein kinase ERK
9	232.5	10.5	367	2	S68680 stress-activated p
10	231.5	10.5	354	2	T45138 protein kinase HOG
11	231	10.5	435	2	S64950 protein kinase (EC
12	230	10.4	298	1	A44878 protein kinase (EC
13	230	10.4	298	2	A41227 protein kinase (EC
14	230	10.4	326	2	S23587 protein kinase (EC
15	230	10.4	373	2	T13024 probable protein k
16	230	10.4	376	2	S40470 mitogen-activated
17	230	10.4	376	2	S40471 mitogen-activated
18	230	10.4	427	2	JC5693 stress-activated p
19	228	10.3	423	2	S43968 p54-alpha-2 stress
20	227	10.3	423	2	S43967 p54-alpha stress-a
21	227	10.3	424	2	S71102 calcium kinase JNK
22	227	10.3	529	1	S71774 calcitum-dependent
23	226	10.2	406	1	KIHUCT phosphotyrase kina
24	225.5	10.2	301	1	S19209 protein kinase (EC
25	225.5	10.2	416	2	A48249 pre-mRNA splicing
26	225	10.2	371	2	S60121 mitogen-activated
27	224.5	10.2	335	2	T23050 hypothetical prote
28	224.5	10.2	342	2	T21098 protein kinase (EC
29	224.5	10.2	371	2	T14915 mitogen-activated

30	224.5	10.2	469	1	S17656 Ca2+/calmodulin-de
31	223.5	10.1	367	2	JC5252 mitogen-activated
32	223.5	10.1	426	2	S43969 p54-beta stress-ac
33	223.5	10.1	464	2	S71104 protein kinase JNK
34	223.5	10.1	678	2	T43539 spindle checkpoint
35	223	10.1	346	1	T48157 protein kinase (EC
36	223	10.1	406	2	D84898 probable mitogen-a
37	223	10.1	427	2	JC5694 stress-activated p
38	223	10.1	575	2	JC7794 lammer kinase homo
39	223	10.1	690	2	T38052 probable protein k
40	222.5	10.1	301	1	A48041 protein kinase (EC
41	222.5	10.1	424	2	A55480 c-Jun amino-termin
42	221	10.0	358	2	S23383 protein kinase (EC
43	220.5	10.0	382	2	JC5531 c-Jun amino-termin
44	220.5	10.0	602	2	S60052 calcitum-dependent
45	220	10.0	371	2	T09622 protein kinase RKK

ALIGNMENTS

RESULT 1
148615
gene KIS protein - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C/Accession: 148615
R/Maucuer, A.; Camonis, J.H.; Sobel, A.
Proc. Natl. Acad. Sci. U.S.A. 92, 3100-3104, 1995
A>Title: Stathmin interaction with a putative kinase and coiled-coil-forming protein dome
A/Reference number: 148282; PMID:95241452; PMID:7724523
A/Accession: 148615
A>Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-200 <RES>
A/Cross-references: UNIPROT:P97343; EMBL:X82320; NID:G791078; PIDN:CAA57763.1; PID:G79107.
A/Genetics:
A:Gene: KIS

Query Match 47.7%; Score 1052; DB 2; Length 200;
Best Local Similarity 99.0%; Pred. No. 8.1e-54;
Matches 198; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY	122	LYSSHQGSMMWIOHCADVLEALFLHHEGVHADLPKRNILMSAENECFKLIDFGLS	181
DB	1	LYSSHQGSMMWIOHCADVLEALFLHHEGVHADLPKRNILMSAENECFKLIDFGLS	60
QY	182	FKRGNDVXYIOTDGYRAPEAEIIONCLAQAGLSPTECTSAVDLWSLGIILMFSGMKL	241
DB	61	FKRGNDVXYIOTDGYRAPEAEIIONCLAQAGLSPTECTSAVDLWSLGIILMFSGMKL	120
QY	242	KRTVMSQEWKANSALIDIFASKAVVNAATPAHYLRDLIKSMHDDPERRIPAEWALCS	301
DB	121	KRTVMSQEWKANSALIDIFASKAVVNAATPAHYLRDLIKSMHDDPERRIPAEWALCS	180
QY	302	PPFSIPFAPHIDVWLPTP	321
DB	181	PPFSIPFAPHIDVWLPTP	200

RESULT 2
T24445
hypothetical protein T04C10.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T24445
R/Burton, J.
submitted to the EMBL Data Library, March 1996
A/Reference number: Z19891
A/Accession: T24445
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-948 <WIL>

A:Cross-references: UNIPROT:O8M0L7; EMBL:Z69885; PDB:CAA3756.1; GSPDB:GNO0028; CESP:T:1
C:Genetic8:
A:Gene: CESP:T04C10.1
A:Map position: X
A:introns: 64/3; 181/3; 214/3; 238/1; 272/3; 331/2; 391/2; 457/1; 726/3; 830/2; 923/3

Query Match 11.3%; Score 248.5; DB 2; Length 948;
Best Local Similarity 25.7%; Pred. No. 5.6e-07;
Matches 104; Conservative 57; Mismatches 113; Indels 131; Gaps 17;

Qy 88 NIVTLGVFTIHFSPPNWPSCRLLLELDVSVSELLLYSHOGSGMMMIQHCARDVLEALA 147
Db 455 NIVTLKGHF-VHRA-----HLCIVLELLSYNYLDLKNYSPFRGVSILNARKFAQQLGKTL 509
Qy 148 FLNHE--GVYHADLKPRNITW-SAENECFKLIDGLSPKESNOQVKTQTDGYAPAEAL 204
Db 510 FLSPPELSIHCDELPEVNLVNNKRSQIRIDGSSCQGTIRLYQIQRHFYSPEVL 569
Qy 205 QNCAAGAGLOSTECTSAVDLSLGIILLEWFSG-----MKL----- 241
Db 570 -----GLAVDTK-----IDWMSLCILVENHTGEPLFAGSSEVDQMKIYEVLGMPKE 618
Qy 242 -----KATVRSQEKANSSAIIIDHPSKAVVNAAI 272
Db 619 MLDIGPKTHKYFDXTEDGIYYCKKTRDGRHT-----YKAGARKLHEILG---VTSQG 669
Qy 273 PA-----YHLRDLIKSMLHDDPSRRIPAEALCSPPFSI--PFAHIEDLV 316
Db 670 PGRRLGEPGHSVEDYSKFKDLIKRLQDFDKAI-----SPYYVRHFPLKQEKERV 722
Qy 317 MLPPVRLNLVLDLDDYLVGNEEYEDVDEKESQKY--GPVSLVLPKENPGRGOVF 374
Db 723 PSQPFVSH-----SNLQQQQOOLYIQPPSSQMSQVMSPEVSGVYV 762
Qy 375 E-----YVAGDSKAQKLLTGRMFDGKF--VVAFTYPLSAVKRY 413
Db 763 EDNGMYROAPGSSANPISVTSFDEGAEVDAGRRRPSAHQNY 807

RESULT 3
A96807
hypochemical protein T3288.5 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: A96807
R:Proteolysis: A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.M.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Hultzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, K.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzalli, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, A.; Xie, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:1130712
A:Accession: A96807
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-720 (#STD>
A:Cross-references: UNIPROT:O9CA22; GB:AE005173; NID:g647532; PDB:AAE08564.1; GSPDB:GNO0028; CESP:T:1
A:Gene: T3288.5
A:Map position: 1

Query Match 10.9%; Score 240.5; DB 2; Length 720;
Best Local Similarity 25.7%; Pred. No. 1.2e-06;
Matches 105; Conservative 57; Mismatches 146; Indels 101; Gaps 17;

Qy 11 EPPRLAEFGMLVQSNLGGSSASVRYRACCGNPSGPPALAQPLPRTGTGAAAGAAE 70
Db 331 DPDLFFKNGKLYRQLGKIGSGSEVHKV---SSDTYALKIKIKLKGVDYTA--- 383

[illegible]

Db 332 -----INDNTKYRVNFRVNVYFVIMRNKMFHSNVL-----NOGESKKEEK 372

RESULT 5

S48879
protein kinase SMK1 (EC 2.7.1.-) - yeast (Saccharomyces cerevisiae)
N/Alternate names: protein YP9499.10; protein YPR054w
C/Species: Saccharomyces cerevisiae
C/Date: 13-Jan-1995 #sequence_revision 10-Feb-1995 #text_change 09-Jul-2004
C/Accession: S48879; SS4076
R/Kristak, L.; Strich, R.; Winters, R.S.; Hall, J.P.; Mallory, M.J.; Kreitzer, D.; Tuan, Genes Dev. 8, 2151-2161, 1994
A/Title: SMK1, a developmentally regulated MAP kinase, is required for spore wall assembly
A/Reference number: S48879; MID:95047362; PMID:7958885
A/Accession: S48879
A/Status: nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 1-388 <RRI>
A/Cross-references: UNIPROT:P41808; EMBL:L35047; NID:G538516; PIDN:AB59325.1; PID:G5385
R/Badcock, K.; Churcher, C.M.
submitted to the EMBL Data Library, May 1995
A/Reference number: S54059
A/Accession: S54076
A/Residues: 1-388 <BAD>
A/Molecule type: DNA
A/Cross-references: EMBL:249219; NID:G805025; PIDN:CAA89172.1; PID:G805035; MIPS:YPR054w
A/Experimental source: strain AB972
C/Genetics:
A/Gene: SGD:SMK1
A/Cross-references: SGD:S0006258; MIPS:YPR054w
A/Map position: 16R
C/Superfamily: kinase-related transforming protein; protein kinase homology
F/Keywords: ATP; phosphotransferase; signal transduction
F/36-306/Domain: protein kinase homology <KIN>
F/44-52/Region: protein kinase ATP-binding motif

Query Match 10.9%; Score 240; DB 2; Length 388;
Best Local Similarity 25.8%; Pred. No. 6,6e-07;
Matches 103; Conservative 61; Mismatches 133; Indels 102; Gaps 19;

QY 21 RLWQVOSRLGSSGSSAVYRVRCCGNPSPGALRQFLPPTGAAASAAEYG----- 72
Db 10 RAINVASNIGAPQQRITFAKERISIPGY--YETIQFLGKAGVGVCSVFKGRSPARIA 67
QY 73 -----PKEK---RAALBQL-----QGHNTVTLVGVTTTHFSNVBSCILLLEL 116
Db 68 VKKISINIFENKEILLKRAIRLKFKNPFKGNKINVLIDETVSSPDGLYC-YOBLIDY 126
QY 117 SVSELLLYSSHQCSMMMIQHCARDVLEALFLHHEGYVHADLKPNNILMSAENECFKLI 176
Db 127 DLAK-VIHSSVQ-LSEFHIKTFYLOLGLKXIHSDVIRDLKPGNIICTL-NGCLKIC 183
QY 177 DFGLS-----FKEGNODVK-----YIQDGYRPAEAEIIONCIAQAGLOSDECTSAVDL 225
Db 184 DFLGANGIHAGFPKCHSTVQPHITNVVATRWYRABEL-----LISNQYKSVDI 233
QY 226 WSGIILLLE-----MESGKLGKHTV-----RSQEW---KANSNAI 257
Db 234 WAGVCLIAEYARAKPVFMGDSMHOIFELIKVLGTDPKDILIKFGTAKANLAKSNNPV 293
QY 258 IDH1-----FASKVAVNAAPAYHLRDLIKSMLHDDPERRIPAEWALCSPPFSIPFAP 310
Db 294 YKKIPSNIPFPASHAIN-----LIESLHMDSIHRILANVQALSHPTLNVKRP 343
QY 311 HIEDLVMLPTPVRLNLNVLDLDYLGNEEYED-VVEDVK 348
Db 344 D-DEPVLQCP-----FDFTYSELSNMSKLDYLVEEVK 377

RESULT 6

T37758
protein kinase skip1 - fission yeast (Schizosaccharomyces pombe)
C/Species: Schizosaccharomyces pombe

C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: T37758
R/Wedler, H.; Wambutt, R.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, December 1998
A/Reference number: 221744
A/Accession: T37758
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-387 <MED>
A/Cross-references: UNIPROT:Q10452; EMBL:AL035064; PIDN:CAA22609.1; GSPDB:GN00066; SPDB:4
A/Experimental source: strain 972h-7; cosmid c1687
C/Genetics:
A/Gene: SPDB:SPAC1687.15
A/Map position: 1
A/Introns: 7/1; 70/3
C/Superfamily: kinase-related transforming protein; protein kinase homology

Query Match 10.8%; Score 238.5; DB 2; Length 387;
Best Local Similarity 24.6%; Pred. No. 8,1e-07;
Matches 78; Conservative 48; Mismatches 112; Indels 79; Gaps 10;

QY 29 LGGSSASVYRVRCCGNPSPGALRQFLPPTGAAASAAEYGRKERRAALBOLQGHN 88
Db 38 VSGSGFVVMQVHLIESDSK--AAIKRVL-----QDKFRKRELQIMRMKIPN 84
QY 89 IYTLGVF-----TIHSPNV---PSRCILLLELDYSVELLYSSHQGS 131
Db 85 IYDLAIAYTTGDNDSDEVIANLVLFMPETIYRASRLYTRQLSMPMLEVLY----- 137
QY 132 MMHICARDVLEALFLHHEGYVHADLKPNNILMSAENECFKLIDFGLS--FKEGNODV 189
Db 138 -----IYQLRSIAYIHASGICHRDIKPKQLLDPENGILKLDGFSAKILVAGEPNV 190
QY 190 KYIQDGYRPAEAEIIONCIAQAGLOSDECTSAVDLMSGIILLFEGS----- 238
Db 191 SYICRRYRABEL-----IFGADYTHAIDIWSTGCVMAELMGHPLPGESEGD 240
QY 239 -----MKLKGTVRSQEWKANSALIDHIFAS-----KAVVNAAPAYHLRDLIKSMLHD 287
Db 241 QLVEIILKILGTSRQIKTMNPNYMEHRRPQIRPPLSRVRSRYPDLAL-DLSKMLQY 299
QY 288 DPSRRIPEMALCSPPF 304
Db 300 TPTDRLTAEMAMCHPFF 316

RESULT 7

S37790
probable serine/threonine-specific protein kinase (EC 2.7.1.-) YKL161c - yeast (Saccharom
N/Alternate names: protein YKL615
C/Species: Saccharomyces cerevisiae
C/Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 16-Aug-2004
C/Accession: S37790; S37991; S44567
R/Vandenbol, M.; Bolle, P.; Dion, C.; Portetelle, D.; Hilger, F.
submitted to the EMBL Data Library, September 1993
A/Description: DNA sequencing of a 36.2 kb fragment located between the FAS1 and LAP4 loc
A/Reference number: S37786
A/Accession: S37790
A/Molecule type: DNA
A/Residues: 1-433 <VAN>
A/Cross-references: UNIPROT:P36005; EMBL:Z26877; NID:G407482; PIDN:CAA81493.1; PID:G40748
A/Experimental source: strain S288C
R/Vandenbol, M.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hilger, F.
submitted to the Protein Sequence Database, March 1994
A/Reference number: S37976
A/Accession: S37991
A/Molecule type: DNA
A/Residues: 1-433 <VA2>
A/Cross-references: EMBL:Z28161; NID:G486280; PIDN:CAA82003.1; PID:G486281; MIPS:YKL161C
A/Experimental source: strain S288C
R/Vandenbol, M.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hilger, F.
Yeast 10, 35-40, 1994
A/Title: DNA sequencing of a 36.2 kb fragment located between the FAS1 and LAP4 loci of c

A:Reference number: S44563
A:Accession: S44567
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-433 <Va3>
A:Cross-references: EMBL:Z26877; NID:G407482; PIDN:CAA81493.1; PID:G407487
A:Experimental source: strain S288C
C:Genetic8:
A:Cross-references: SGD:S0001644
A:Map position: 11L
A:Note: YKL161c
C:Superfamily: protein kinase homology
C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F:21-287/Domain: protein kinase homology <KIN>
F:153/Active site: Asp #status predicted

Query Match	10.6%;	Score 234.5;	DB 2;	Length 433;
Best Local Similarity	29.5%;	Pred. No. 1.5e-06;		
Matches 89;	Conservative 45;	Mismatches 97;	Indels 71;	Gaps 16;

```

Oy 75 KERALBEOLOGRNRNITVLXGVFTIHSPVPRBCLL--ELDVSSELLLXSHQCGSM 132
Db 71 RELKLRLHRLRGPNITVWLFDT-DIVFYPGALNGVLYKEELMECDISQIT--RSEQRLED 127
Oy 133 WMIQHCARDVLEALAFILHHEGYVHADIKRPNILMSANEC-FKLIDFGIS-----FKEG 185
Db 128 AHFQSFYQIICALKXIHSAVHCDIKRNLL--VNSDCQIKICNFGLSCSYSEHNKYN 185
Oy 186 NQDYK-YIQTDCGYRARELONCLAOAGIQSDTCTSAVDLWSLGIILLE-----MFSG 238
Db 186 DGLIGYITSIWYKAPET-----LLNQECTKAVDINSTGCTIABLGRKPMFEG 235
Oy 239 M---KLKHYR-----SQEWKANSALIDHIF-----ASKVYVNAALPAYNL 277
Db 236 KDYVDHLNHLIQLIGTPEETLOE--IASQKYVNYFCGNIIGRSFSEILGAMP--EA 291
Oy 278 RDLIKSMHDDPSRRIRPAMALCSPPFSI-----PFA-----PHIEDLWMLPTFV 322
Db 292 LELIKMLEPFPKKRIIVEDALEHPYLSWMHDIDEEFSCQKTRFPEFEHIESMAELGNEV 351
Oy 323 LR 324
Db 352 IK 353

```

RESULT 8
A:Accession: A56492
Protein kinase ERK2 (EC 2.7.1.-) - slime mold (*Dictyostelium discoideum*)
C:Species: *Dictyostelium discoideum*
C:Date: 19-Oct-1995 #sequence 19-Oct-1995 #text_change 09-Jul-2004
A:Accession: A56492
R:Seiwall, J.E.; Kuapa, A.; Shaulsky, G.; Ecker, M.; Maeda, M.; Gaskins, C.; Firtel, R.A.
J. Cell Biol. 128, 405-413, 1995
A:Title: A MAP kinase necessary for receptor-mediated activation of adenylyl cyclase in
A:Reference number: A56492; PMID:95146547; PMID:7844154
A:Accession: A56492
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-369 <SEG>
A:Cross-references: UNIPROT:O7M445; GB:I133043
C:Superfamily: Kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphotransferase
E:12-273/Domain: protein kinase homology <KIN>
E:20-28/Region: protein kinase ATP-binding motif

Query Match	10.6%;	Score 233;	DB 2;	Length 369;
Best Local Similarity	26.9%;	Pred. No. 1.6e-06;		
Matches	91;	Conservative	52;	Mismatches 109;
			Indels	86;
			Gaps	17;

QY 66 ASAAYGFRERAALEQLOGRNIVTLXGVFTIHFSNVPSRCILLETLDVSVSEL---- 121

Db 52 ATDQRTER-EIMFLOELHGHENITKLVNK--ADNDRDIYLVVEHMETDLDHAVRAK 107

```

Qy 122 L1TSSHQGSGMMMI0HCARDVLEALAF1HNGUYHAD,KPKNIT1MSXNCF-KLIDPG1 180
Db 108 ILEIHK-----QYTIYQLLKALKYHNSANVLHBDIKPSNLT--INSECLVKVADFG1 158
Qy 181 -----SPKGNODV-KYIOTDGYRARELONCLAGLSQSDTCTSAVDMISGII 231
Db 159 ARSITSLESIAEANPYLEYATRYARPAEI-----LGSITRYTKGVDMISGCI 208
Qy 232 ILE-----MFGS-----MKLKATVISOEWKANSALIHIPASKAVVN----- 269
Db 209 LGBLLBEXKAMFPGNSMTNQDLIIEVTRPSAEDIEMAKSPFACTWLES,PPSNPSLS0 268
Qy 270 ----A1PAPHRLDLIKSM1HDDPSRR1PAEMALCSPP--FS1PF-AH1IEDVWMLPT- 320
Db 269 MYSASVVDL--DLIKKLSOFNPKRITTAEBALAHPPVYQFHNPAEHPHDIRIKISID 325
Qy 321 ----PVLRLNLVLDYDYLGNEESEYEDVAVEBKEEC0 353
Db 326 DGQKPIAEYRNL-----YNDILIKKKKBERKK 353

```

```
QY      321 -----PVLRLNLVDDDYLGNEEYEDVEDVVEEQCK 355
          |::|
D6      326 DGGKEPIAEYRNL-----YNDIKKKKEERKK 355
```

RESULT 9
 S68680 stress-activated protein kinase (EC 2.7.1.-) 3 - rat
 C.Species: *Rattus norvegicus* (Norway rat)
 C.Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
 C.Accession: S68680
 R.McCortens, S.; Craxton, M.; Goedert, M.
 FEBS Lett. 383, 273-276, 1996
 A.Title: SAP kinase-3, a new member of the family of mammalian stress-activated protein
 A.Reference number: S68680; MUID:96198618; PMID:8925912
 A.Accession: S68680
 A.Status: Preliminary
 A.Molecule type: mRNA
 A.Residues: 1-367 <MR>
 A.Cross-references: UNIPROT:O65338; EMBL:X96488; NID:G1262402; PIDN:CAA65342.1; PID:G1262402
 C.Superfamily: kinase-related transforming protein; protein kinase homology
 C.Keywords: ATP, phosphotransferase
 P.235-280/Domain: protein kinase ATP-binding motif
 P.733-41/Region: protein kinase ATP-binding motif

Query March	10.5%	Score 232.5;	DB 2,	Length 367;
Best Local Similarity	22.6%	Pred. No. 1.7e-06;		
Matches 95; Conservative	68;	Mismatches 127;	Indels 131;	Gaps 19

```

OY 10 AEBPRFLEAAGR-----LMOYOS-----RIGSSSSAAYRVACCGNPGSPPE-----ALR 54
Db 2 SSSPPARKGFYRQEWYTKTAMEYRAVAYODLOPYGSAIYAV-----CSANDSRGNKKAIK 56
OY 55 QPLPBGTTGAASAAYEYGRKERAALEOLQGHRIVTLVYGVFTTHFSPNVPSSRC----- 1089
Db 57 KLYRPFQSELPFAKRA-----YRELRLILKKMR-HENVIIGLDIVFT-----PDETLDDFTD 104
OY 109 --LLELBDVSYSELLLYSHOGCSMMNIOHCARDVLEALAFCHHEGVHADLCKRILW 166
Db 105 FYLWPMFGMTDGLKLM--KHETTSIEDRIQELVYQMKLGKLYHAAAGVITHDLKPGVL-- 159
OY 167 SAENEC-FKLIIDGYSFKFEGNODVYKIIOTDGYRAPEALONCLQAQIQSDETECTSAVDL 225
Db 160 ANVEDDEKLIDPGLARQADSEMTGYVTRMYRAPEVL-----XMMRTQYVDI 209
OY 226 WSLGIIILEMFGSMKILKHTVRSQEWKANKSSAIIIDHIFASKAVVNA----- 270
Db 210 WSYGCIAMAEWITOKIL-----FKGN-----DHLDLKEIMKVTGHPPEFOYQLQA 256
OY 271 -----ALPANHLR-----LKSXTHLDBPSRRIRPAEMALCSPPFSIIPFA 309
Db 257 EAKNYMEGPELEKKODFASVLTNAPSPAVNLLEKXLYADABQRYTAEBALAHPEYF----- 312
OY 310 PHIEDLVMLPTPLRLINVLDLDDDYLGNEEYEDVYEDVK-----EEOCK--YGPVVSLVVPK 364
Db 313 -----SIRDTEDEPRAKQKODSDDDVDRLTEEMKRVYTKGVUSFKEPR 355

```

Qy 365 E 365
Db 356 Q 356

RESULT 10

T45138
protein kinase ekpl [imported] - fission yeast (Schizosaccharomyces pombe)
C/Species: Schizosaccharomyces pombe
C/Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C/Accession: T45138
R/Player, S.E.; Feoktistova, A.; Burke, J.D.; Woodgett, J.R.; Gould, K.L.
Mol. Cell. Biol. 16, 179-191, 1996
A/Title: Schizosaccharomyces pombe ekpl+ encodes a protein kinase related to mammalian g
A/Reference number: 2252927, MIMD:96104567, PMID:8524294
A/Accession: T45138
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-354 <PLY>
A/Cross-references: EMBL:L29449; NID:g1914882; PIDD:AA51081.1; PID:g1914883
C/Genetics:
A/Genes: ekpl
A/Intons: 6/3
C/Function:
A/Pathway: cells w
A/Note: cells with deletions in ekpl are sensitive to heat shock and exhibit defects in
C/Keywords: phosphoprotein, phosphotransferase
F/335/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match 10.5%; Score 231.5; DB 2; Length 354;

Best Local Similarity 24.3%; Pred. No. 1.8e-06;

Matches 77; Conservative 48; Mismatches 113; Indels 79; Gaps 10;

Qy 29 LSGSSASVYVRCCGPPGALROFLPFGTTGAASAAYGFKERAALEFOLQGHNN 88
Db 38 VSGSGVGVVQVALLISDSK-AAIKRVL-----QDKRFKRELOIRIMNDPN 84
Qy 89 IYLLYGVF-----TIHSPNV---PSRCLLELLDVSVSELLYSSHQGS 131
Db 85 IYLLIYVTTGDNSEVYINLVLEMPETTYRASRLYTRQKLSMPLEVKY----- 137
Qy 132 MMNIQHCADVLEALAFHHEGYVHADLKPRNIIIMSAENECFLIDFGLS--FKEGNOY 189
Db 138 -----IYQLASLAVIHASGICHRDIKQNLILPENCILKCPFGSKALIVAEPNV 190
Qy 190 KYIQDGYRAPEALQNCIAQAGLQSDPECTSAVDLMSLIGILLMFSG----- 238
Db 191 SYICSRYYRAPEL-----IFGATDYTHAIDIMSTGCVAAELMLGHPLPFGSGID 240
Qy 239 -----MKLKITVSQEMKANSALIDHIFAS-----KAVNNAAIRAHRLDIKSMND 287
Db 241 QLVETIKILGTPSREDIKTNNPNYMERFPQIRPPLSRVSRSPVLDAI-DLISKLOY 299
Qy 288 DPSRIIPAEWALCSPPF 304
Db 300 TPTDRILTAEMAKCHPFF 316

RESULT 11

S64950
protein kinase HOG1 (EC 2.7.1.-), mitogen-activated - yeast (Saccharomyces cerevisiae)
N/Altemate names: protein l2931; protein YLR113w
C/Species: Saccharomyces cerevisiae
C/Date: 01-Aug-1995 #sequence_revision 24-May-1996 #text_change 16-Aug-2004
C/Accession: S64950; S42145; S69400
R/Verhaaselt, P.; Voelt, M.; Volckaert, G.
submitted to the Protein Sequence Database, May 1996
A/Reference number: S6443
A/Accession: S64950
A/Molecule type: DNA
A/Residues: 1-435 <VER>
A/Cross-references: UNIPROT:P32485; EMBL:Z73285; NID:g1360507; PIDD:CAA97680.1; PID:e245

A/Experimental source: strain S288C
R/Brewster, J.L.; de Valoir, T.; Dwyer, N.D.; Winter, B.; Gustin, M.C.
Science 259, 1760-1763, 1993
A/Title: An osmosensing signal transduction pathway in yeast.
A/Reference number: S42145, MIMD:93206121, PMID:7681220
A/Accession: S42145
A/Status: nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 1-408 'G08CSCCK' <BRE>

A/Cross-references: EMBL:L06279; NID:g295610; PIDD:AAA34680.1; PID:g295611
R/Verhaaselt, P.; Volckaert, G.
submitted to the EMBL Data Library, September 1995
A/Reference number: S69393
A/Accession: S69400
A/Molecule type: DNA
A/Residues: 1-435 <VER>

A/Cross-references: EMBL:X89514; NID:g1297019; PIDD:CAA61691.1; PID:e198746; PID:g129702;
C/Genetics:
A/Genes: SGD:HOG1; SSK3
A/Cross-references: SGD:S0004103; MIPS:YLR113w
A/Map position: 12R
C/Superfamily: protein kinase homology
C/Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase; signal tr
F/21-271/Domain: protein kinase homology <KIN>
F/29-37/Region: protein kinase ATP-binding motif
F/144/Active site: Asp #status predicted

Query Match 10.5%; Score 231; DB 2; Length 435;

Best Local Similarity 25.5%; Pred. No. 2.5e-06;

Matches 97; Conservative 54; Mismatches 140; Indels 90; Gaps 18;

Qy 17 EAFGRLMVOYSR-----LSGSSASVYVRCCGNP--GSPGALROFLPFGTTGAASA 68
Db 11 QIFGVYFETIRNYNDLNPVGMARGLV---CSAIDTLTSGVALLKIMKPESTAVLAAR 66
Qy 69 AEYGFKEKPALEQOGHENVLYGVFTIHSPNVPSRCLLELLDVSVSELLYSSHQ 128
Db 67 T-----YREIKLKHRL-HENILCLQDIF---LSP-LIEDYFVTELOQTDILRLT---QTR 114
Qy 129 GCSMMIQCADVLEALAFHHEGYVHADLKPRNIIIMSAENECFLIDFGSPFEGNO 188
Db 115 PLEKQFOYFLQIRGLKYHAGVHRDLKPSNILL-ENCDDKICDGLARIQDPQM 173
Qy 189 VKYIQDGYRAPEALQNCIAQAGLQSDPECTSAVDLMSLIGILLMFSGMKL----KAT 244
Db 174 TGYSTRYRAPELML-----TWQKIVY---VDINSAGCIFAMIRGKLPFGKDH- 222
Qy 245 VRSQEMKANSALIDHIFAS--KAVNNAAIRAHRL----- 278
Db 223 -----VHQFSIITDLGSPKDVINTICSENTLKFTVSLPHRDPFPSRFTVEPDA 275
Qy 279 -DLISKMLDDPSRKIPAEWALCSPPFSPFPAHLEDVLM-----LPTPVLR 325
Db 276 VDLLEKMLVDFDKKITTADALANP-YSAFYHDPDEPVADAKFPMHFNDAQLPVDYWRV 334
Qy 326 L---NYLDDVYLGNEEYEDV 343
Db 335 MMYSLELDPHKIGSGDQIDT 355

RESULT 12

A44878
protein kinase (EC 2.7.1.37) cdk2 [validated] - goldfish
C/Species: Carassius auratus (goldfish)
C/Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C/Accession: A44878
R/Hirai, T.; Yamashita, M.; Yoshikuni, M.; Tokumoto, T.; Kajitani, H.; Sakai, N.; Nagahama
Dev. Biol. 152, 113-120, 1992
A/Title: Isolation and characterization of goldfish cdk2, a cognate variant of the cell c
A/Reference number: A44878; MIMD:92331802; PMID:1339336
A/Accession: A44878
A/Status: preliminary
A/Molecule type: mRNA

	Matches	87; Conservative	42; Mismatches	124; Indels	64; Gaps	10;
CY	23	WQVQRSLGSSGSAYTVRVCCNPPSPGALRQFLPRGGTGAASAAGYGFKEKPALEAQ	82			
Dd	13	YECVAIEIGEGAVGVKKFKARDLNGGRFVALKVRVQTGEEMPLSTI-----REVAVLRLH	67			
CY	83	LQ--GHRIIVTLVGFTIHFSFNVSRCLLLELDVSVSEILLYSHGCSMMIMQHCAR	140			
Dd	68	LETPEHPVVNRLEFDVCTVSRIDRETETKLTLVEFHVODLTYYLDKYPERPVPTEITKDMMF	127			
CY	141	DYLEALTAELTHHEGYHADLKPRNIIMSASNECFKILIDFGI-----SFEKGNDPVKIQTGD	196			
Dd	128	QLRLRDPLTHSRVYRHDLKPQNIIIVTSSGQ--IKLADYGARITYSFQALNTSV--VTLM	184			
CY	197	YRAPEALQNCIAOGLQSDBTECTSAYVDLWSGIILILEFSGMKLKHTVRSQEWKANSNA	256			
Dd	185	YRAPEVLLQSSYA-----TPVDLMSVGCIFAEWFRKKPL-----FRGSS	223			
CY	257	IIDHIFASKAVV-----NAALP--AYNL-----RDLIKSMTLHD	287			
Dd	224	DVDQLGKITLVYIGLPGEEDWRPDVALPROAFHSKSAPIEKEFTVIDELGKDLLKCLTF	283			
CY	288	DPSRRIPAEMLACSPFF	304			
Dd	284	NPAKTSISAYSLSHPYF	300			

```
QY      344 VEDVKE-----ECQKYGPPVSL 360
          |::|| |::|:|:
Db      353 EEI KELVWLESVKFNPLPSI 373
```

Search completed: November 29, 2004, 11:42:41
Job time : 27 secs

RESULT 15

T13024

probable protein kinase (EC 2.7.1.-) F8L21.120 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 21-Jan-2000

C:Accession: T13024

R:Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; Mew

submitted to the Protein Sequence Database, July 1999

A:Reference number: Z1587

A:Accession: T13024

A:Molecule type: DNA

A:Residues: 1-373 <BEV>

A:Cross-references: EMBL:AL096882, GSPDB:GN00062, ATSP:F8L21.120

A:Experimental source: cultivar Columbia; BAC clone F8L21

C:Genetics:

A:Gene: ATSP:F8L21.120

A:Map position: 4

A:Mutrons: 58/1; 98/3; 144/3; 255/3; 317/1

C:Superfamily: kinase-related transforming protein; protein kinase homology

C:Keywords: phosphotransferase; protein kinase

C:Keywords: /40-295/Domain: protein kinase homology <KIN>

Query Match	10.4%	Score 230	DB 2	Length 373
Best Local Similarity	25.9%	Pred. NO. 2.4e-06		
Matches 83, Conservative	59	Mismatches 111	Indels 68	Gaps 14

```

Qy      75  KERALBLOOGRNIVLTJYGVFTTHFSNPNVPRCLLELTVSVSELLYSHQCSMM 134
Db      86  REIKLRHLE-HENVVVKIDIRPKKDFVDVYIVFELMDTDHQII--KSNQSLNDH 142
Qy      135  IOHCARDVLEALAFIHHGYVHADLKPRNIIIMSAENECFKLIDFEL--SFKGNODVYI 192
Db      143  COYFLYQIIRGKLYHSANVLRHDKPSNLNLS--NCPLKIDFELARTSTETRYMTEYV 201
Qy      193  QTDGVRAPAEALONLQAQLSDPDECSAVDMLSGIILE-----MFSQMKLHYRV 246
Db      202  VTRWYRAPEL-----LINSEYTSALIDVMSVGCIPAEIMTRREPIFPKQDYHQK 251
Qy      247  -----SOEW--KANSSAIIHH-----PASK--AVYNAAI PAYTHDLIKS 283
Db      252  LITELIGSPDGASLEFLRSANARKYVKELPKPRQNFSAKSPSNMSTAI-----DLLEK 305
Qy      284  MLHDDPSRRIPAEMLGSPFESIPPAPIHEDVYMLPTVRLRLNVLDDDYLANEEYEDV 343
Db      306  MLVGPVPRIVRIVEEALCYPYLSA-----LHDLNDEP-----VCSNHFSPHFHPDPST 352

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2004, 11:26:47 ; Search time 122.5 Seconds
(without alignments)
1968.016 Million cell updates/sec

Title: US-10-798-532-4

Perfect score: 2206
Sequence: 1 MAGSGCAGCABPPRFLEARG.....VATFVPLSAVKRGVLYQTLL 419

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2195	99.5	419	1 KIST_HUMAN	Q81a81 homo sapien
2	2183	99.0	419	1 KIST_RAT	Q63785 rattus norv
3	2182	98.9	419	2 AABH5732	AABH5732 mus muscu
4	2182	98.9	419	2 AAO13515	AAO13515 mus muscu
5	2177	98.7	419	1 KIST_MOUSE	P97343 mus musculu
6	1711	77.6	330	2 Q8C0N6	Q8C0N6 mus musculu
7	254	11.5	505	2 Q6CFX3	Q6CFX3 yarrowia li
8	248.5	11.3	882	2 Q8W0L7	Q8W0L7 caenorhabdi
9	242	11.0	433	2 Q84SN3	Q84SN3 oryza sativ
10	240.5	10.9	395	2 Q6GM90	Q6GM90 xenopus lae
11	240.5	10.9	720	2 Q9CA22	Q9CA22 arabidopsis
12	240.5	10.9	765	2 Q25758	Q25758 plasmodium
13	240.5	10.9	777	2 Q84VX4	Q84VX4 arabidopsis
14	240.5	10.9	826	2 Q94656	Q94656 plasmodium
15	240.5	10.9	914	2 Q81IF0	Q81IF0 plasmodium
16	240	10.9	368	1 SMK1_YEAST	P41808 saccharomyc
17	239.5	10.9	395	2 Q8AYR3	Q8AYR3 xenopus lae
18	239.5	10.9	395	1 AAH70745	AAH70745 xenopus l
19	238.5	10.8	387	1 SKP1_SCHPO	Q10452 schizosacch
20	237.5	10.8	360	2 Q6S267	Q6S267 lytechinus
21	237.5	10.8	360	2 AAR17088	AAR17088 lytechinu
22	235.5	10.7	434	2 Q75B85	Q75B85 ashbya goss
23	235.5	10.7	434	2 AAS51605	AAS51605 ashbya go
24	235	10.7	298	1 CDR2_MESAU	P48963 mesocricetu
25	234.5	10.6	392	2 Q6ZD93	Q6ZD93 oryza sativ
26	234.5	10.6	392	2 BAC9508	BAC9508 oryza sat
27	234.5	10.6	433	1 KKO1_YEAST	P36005 saccharomyc
28	234	10.6	376	1 MPK4_ARATH	Q39024 arabidopsis
29	234	10.6	376	1 AAM66070	AAM66070 arabidops
30	233.5	10.6	433	2 Q6B829	Q6B829 paramecium
31	233.5	10.6	1233	2 Q6CAC2	Q6CAC2 yarrowia li

32	233	10.6	369	2 Q7M445	Q7M445 dictyosteli
33	232.5	10.5	367	1 MK12_RAT	Q63538 rattus norv
34	232	10.5	568	2 Q9HB76	Q9HB76 homo sapien
35	232	10.5	588	2 Q9HB77	Q9HB77 homo sapien
36	232	10.5	588	2 AAT06103	AAT06103 homo sapi
37	231.5	10.5	367	1 MK12_MOUSE	Q08911 mus musculu
38	231.5	10.5	464	1 MK10_RAT	P49187 rattus norv
39	231.5	10.5	542	2 Q61587	Q61587 oryza sativ
40	231	10.5	298	1 CDK2_CRIGR	Q55076 cricetulus
41	231	10.5	298	1 CDK2_RAT	Q63699 rattus norv
42	231	10.5	298	2 Q6F751	Q6F751 rattus norv
43	231	10.5	298	2 AAB61832	AAB61832 rattus n
44	231	10.5	361	1 JNK_SUBDO	Q966y3 subdites d
45	231	10.5	435	1 HOG1_YEAST	P32485 saccharomyc

ALIGNMENTS

RESULT 1
KIST_HUMAN
ID KIST_HUMAN STANDARD; PRT; 419 AA.
AC Q8TAS1; Q96C22;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Serine/threonine-protein kinase Kist (EC 2.7.1.37) (Kinase interacting with statmin).
GN Name=KIST; Synonyms=KIS;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX PubMed=12782333;
RA Bieche I., Manceau V., Curni P.A., Laurendeau I., Lachkar S., Leroy X., Vidand D., Sobel A., Maucuer A.;
RT "Quantitative RT-PCR reveals a ubiquitous but preferentially neural expression of the KIS gene in rat and human."
RL Brain Res. Mol. Brain Res. 114:55-64(2003).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Eye, and Skin;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D., Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D., Bosa S.A., Loquellano N.A., Peters G.J., Malak J.A., Gunaratne P.H., Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.D., Uesdin T.B., Toshiyuki S., Carninci P., Prange C., Bosa S.A., McEwan P.J., McKernan K.J., Malak J.A., Gunaratne P.H., Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E., Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Possesses kinase activity. May be involved in trafficking and/or processing of RNA (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- SUBUNIT: Interacts with statmin (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;

```

CC      IsoId=Q8TRAS1-1; Sequence=Displayed;
CC      Name=2;
CC      Note=No experimental confirmation available;
CC      -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC      -1- SIMILARITY: Contains 1 RNA recognition motif (RNM) domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AC156197; CAD60192.1; -
DR      EMBL; BC014917; AAH14917.1; -
DR      EMBL; BC026046; AAH26046.1; -
DR      HSSP; P26368; 100P.
DR      GO; GO:0005737; C:cytoplasm; ISS.
DR      GO; GO:0005634; C:nucleus; ISS.
DR      GO; GO:0005524; P:ATP binding; ISS.
DR      GO; GO:0005515; F:protein binding; ISS.
DR      GO; GO:0004674; F:protein serine/threonine kinase activity; ISS.
DR      GO; GO:0003723; F:RNA binding; ISS.
DR      GO; GO:0018105; P:peptidyl-serine phosphorylation; ISS.
DR      InterPro; IPR011009; Kinase like.
DR      InterPro; IPR000719; Prot kinase.
DR      InterPro; IPR000504; RNA Rec mot.
DR      InterPro; IPR008271; Ser_Thr_pkin_AS.
DR      Pfam; PF00069; PKinase; 1.
DR      Pfam; PF00076; RNM_1; 1.
DR      ProDom; PD000001; Prot_kinase; 1.
DR      SMART; SM00360; RNM; 1.
DR      PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR      PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR      PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
DR      PROSITE; PS00102; RNM; 1.
DR      PROSITE; PS00030; RNM_RNP_1; FALSE_NEG.
KW      Alternative splicing; ATP-binding; Nuclear protein; RNA-binding;
KW      Serine/threonine-protein kinase; Transferase.
FT      DOMAIN 23..304
FT      NP_BIND 324..406
FT      NP_BIND 29..37
FT      BINDING 54..54
FT      ACT_SITE 158..158
FT      VARSPIC 342..344
FT      VARSPIC 345
FT      VARSPIC 419
FT      SEQUENCE 419 AA; 46546 MW; 903E982BE12A8CF8 CRC64;
SQ
Query Match      99.5%; Score 2195; DB 1; Length 419;
Best Local Similarity 99.5%; Pred. No. 2.4e-158;
Matches 417; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

DB      241 LKHTVRSQEWKANSALIDHIFASAKAVVNAIPAYHLRDLIKSMHDDPSRRIPAEMLC 300
QY      301 SPFSIPFAPHIEDVLMPTPLRLNVLDDDYLCNEEYEDVDVDECKQYPPVSL 360
DB      301 SPFSIPFAPHIEDVLMPTPLRLNVLDDDYLCNEEYEDVDVDECKQYPPVSL 360
QY      361 LVPKENPGRGQFVEVYANAGDSKAAQKLTGRMPDGKVFVATFPLSAVKRGYLYOTLL 419
DB      361 LVPKENPGRGQFVEVYANAGDSKAAQKLTGRMPDGKVFVATFPLSAVKRGYLYOTLL 419

RESULT 2
KIST RAT      STANDARD;      PRT;      419 AA.
ID      KIST RAT
AC      063285;
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DE      05-JUL-2004 (Rel. 44, Last annotation update)
DE      Serine/threonine-protein kinase Kist (EC 2.7.1.37) (kinase interacting
DE      with strachin) (PAM COOH-terminal interactor protein 2) (P-CIP2).
GN      Name=Kist; Synonyms=Kis;
OS      Rattus norvegicus (Rat);
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_TaxId=10116;
RN      [1]
RP      SEQUENCE FROM N.A., FUNCTION, SUBUNIT, SUBCELLULAR LOCATION, TISSUE
RP      SPECIFICITY, AND MUTAGENESIS OF LYS-54.
RC      TISSUE=striatum;
RX      MEDLINE=97435279; PubMed=9287318;
RA      Maucuer A., Ozon S., Manceau V., Gavet O., Lawler S., Curmi P.,
RA      Sobel A.;
RT      "Kis is a protein kinase with an RNA recognition motif.";
RL      J. Biol. Chem. 272:23151-23156(1997).
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Caldwell B.D., Darlington D.N., Penzes P., Johnson R.C., Eipper B.A.,
RA      Mains R.E.;
RL      Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE OF 28-419 FROM N.A.
RC      TISSUE=hippocampus;
RX      MEDLINE=97067094; PubMed=8910496;
RA      Alam R., Caldwell B.D., Johnson R.C., Darlington D.N., Mains R.E.,
RA      Eipper B.A.;
RT      "Novel proteins that interact with the COOH-terminal cytosolic routing
RT      determinants of an integral membrane peptide-processing enzyme.";
RL      J. Biol. Chem. 271:28636-28640(1996).
CC      -1- FUNCTION: Possesses kinase activity. May be involved in
CC      trafficking and/or processing of RNA.
CC      -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphorylated protein.
CC      -1- SUBUNIT: Interacts with strachin (By similarity).
CC      -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.
CC      -1- TISSUE SPECIFICITY: In the embryo, preferentially expressed in the
CC      developing nervous system.
CC      -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC      -1- SIMILARITY: Contains 1 RNA recognition motif (RNM) domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X98374; CAA67021.1; -
DR      EMBL; U70372; AAC53031.2; -
DR      HSSP; P26368; 100P.
DR      RGD; 2968; Kist.
DR      GO; GO:0005737; C:cytoplasm; IDA.
DR      GO; GO:0005634; C:nucleus; IDA.

```

DR GO; GO:0005524; F:ATP binding; IC.
DR GO; GO:0005515; F:protein binding; ISS.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IDA.
DR GO; GO:0003723; F:RNA binding; NAS.
DR GO; GO:0018105; P:peptidyl-serine phosphorylation; IDA.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR000719; Prot_Kinase.
DR InterPro; IPR005054; RNA_rec_mot.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00069; Pkinase; I.
DR Pfam; PF00076; RRM_1; 1.
DR ProDom; PD000001; Prot_Kinase; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; FALSE_NEG.
DR PROSITE; PS50011; PROTEIN KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNF_1; FALSE_NEG.
KW ATP-binding; Nuclear protein; RNA-binding;
Serine/threonine-protein kinase; Transferase.
FT DOMAIN 23 304 Protein kinase.
FT NP_BIND 324 406 RNA-binding (RRM).
FT BINDING 29 37 ATP (By similarity).
FT ACT_SITE 141 141 ATP.
FT ACT_SITE 158 158 Proton acceptor (By similarity).
FT MUTAGEN 54 54 K->R: Loss of activity.
SQ SEQUENCE 419 AA; 46505 MW; D2F9B7BF8080DF3 CRC64;

Query Match 99.0%; Score 2183; DB 1; Length 419;
Best Local Similarity 98.8%; Pred. No. 1.9e-157;
Matches 414; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAGSGCAGAAEPPEPFLAEGRLMVOVSRLGSSGSASVYRVCCGPGSPGALROFLPG 60
DB 1 MAGSGCAGAAEPPEPFLAEGRLMVOVSRLGSSGSASVYRVCCGPGSPGALROFLPG 60
QY 61 TTGAASAABEYGFRRKRALBOLQGHRIYTLGYFTIHFPSPVPSRCILLELDVSVSE 120
DB 61 TTGAASAABEYGFRRKRALBOLQGHRIYTLGYFTIHFPSPVPSRCILLELDVSVSE 120
QY 121 LLLYSSHGSCMMWMIQHCAVDYLEALAFLLHGGYVHADLKPRNIIWASAENECFKLIDFGL 180
DB 121 LLLYSSHGSCMMWMIQHCAVDYLEALAFLLHGGYVHADLKPRNIIWASAENECFKLIDFGL 180
QY 121 LLLYSSHGSCMMWMIQHCAVDYLEALAFLLHGGYVHADLKPRNIIWASAENECFKLIDFGL 180
DB 121 LLLYSSHGSCMMWMIQHCAVDYLEALAFLLHGGYVHADLKPRNIIWASAENECFKLIDFGL 180
QY 181 SFKEGNQDVKYIQTGYRPAPEALONCLAQGLQSDTECTSAVDLMSIGIILLEMFSGMK 240
DB 181 SFKEGNQDVKYIQTGYRPAPEALONCLAQGLQSDTECTSAVDLMSIGIILLEMFSGMK 240
QY 241 LKHTVRSOEWKANSALIDHIFASKAVVNAAPAYHLRDLIKSMHDDPSRRIPEAMALC 300
DB 241 LKHTVRSOEWKANSALIDHIFASKAVVNAAPAYHLRDLIKSMHDDPSRRIPEAMALC 300
QY 301 SPFFSIIPAPHIEDLVMLPTPVRLRLNVLDDDYLGNEEYEDVVDVKECKQYGPVSL 360
DB 301 SPFFSIIPAPHIEDLVMLPTPVRLRLNVLDDDYLGNEEYEDVVDVKECKQYGPVSL 360
QY 361 LVPKENPGRGVFVEYANAGDSKAOKLLTGMPFGKVVATFYPLSAVKGYLQYTL 419
DB 361 LVPKENPGRGVFVEYANAGDSKAOKLLTGMPFGKVVATFYPLSAVKGYLQYTL 419

RESULT 3
AAH58732
ID AAH58732 PRELIMINARY; PRT; 419 AA.
AC AAH58732;
DT 02-MAR-2004 (Tremblrel. 27, Created)
DT 02-MAR-2004 (Tremblrel. 27, Last sequence update)
DT 02-MAR-2004 (Tremblrel. 27, Last annotation update)
DE Kinase interacting with leukemia-associated gene (Statmin).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OK NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RC MEDLINE=22386257; PubMed=12477932;
RA Strusberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalón D.K., Mizny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kertman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butcherfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallus D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strusberg R.;
RL Submitted (Sep-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC058732; AAH58732.1; -.
KW Kinase.
SQ SEQUENCE 419 AA; 46489 MW; 3BD3C06A59E22540 CRC64;

Query Match 98.9%; Score 2182; DB 2; Length 419;
Best Local Similarity 98.8%; Pred. No. 2.3e-157;
Matches 414; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAGSGCAGAAEPPEPFLAEGRLMVOVSRLGSSGSASVYRVCCGPGSPGALROFLPG 60
DB 1 MAGSGCAGAAEPPEPFLAEGRLMVOVSRLGSSGSASVYRVCCGPGSPGALROFLPG 60
QY 61 TTGAASAABEYGFRRKRALBOLQGHRIYTLGYFTIHFPSPVPSRCILLELDVSVSE 120
DB 61 TTGAASAABEYGFRRKRALBOLQGHRIYTLGYFTIHFPSPVPSRCILLELDVSVSE 120
QY 61 TTGAASAABEYGFRRKRALBOLQGHRIYTLGYFTIHFPSPVPSRCILLELDVSVSE 120
DB 61 TTGAASAABEYGFRRKRALBOLQGHRIYTLGYFTIHFPSPVPSRCILLELDVSVSE 120
QY 121 LLLYSSHGSCMMWMIQHCAVDYLEALAFLLHGGYVHADLKPRNIIWASAENECFKLIDFGL 180
DB 121 LLLYSSHGSCMMWMIQHCAVDYLEALAFLLHGGYVHADLKPRNIIWASAENECFKLIDFGL 180
QY 121 LLLYSSHGSCMMWMIQHCAVDYLEALAFLLHGGYVHADLKPRNIIWASAENECFKLIDFGL 180
DB 121 LLLYSSHGSCMMWMIQHCAVDYLEALAFLLHGGYVHADLKPRNIIWASAENECFKLIDFGL 180
QY 181 SFKEGNQDVKYIQTGYRPAPEALONCLAQGLQSDTECTSAVDLMSIGIILLEMFSGMK 240
DB 181 SFKEGNQDVKYIQTGYRPAPEALONCLAQGLQSDTECTSAVDLMSIGIILLEMFSGMK 240
QY 241 LKHTVRSOEWKANSALIDHIFASKAVVNAAPAYHLRDLIKSMHDDPSRRIPEAMALC 300
DB 241 LKHTVRSOEWKANSALIDHIFASKAVVNAAPAYHLRDLIKSMHDDPSRRIPEAMALC 300
QY 301 SPFFSIIPAPHIEDLVMLPTPVRLRLNVLDDDYLGNEEYEDVVDVKECKQYGPVSL 360
DB 301 SPFFSIIPAPHIEDLVMLPTPVRLRLNVLDDDYLGNEEYEDVVDVKECKQYGPVSL 360
QY 361 LVPKENPGRGVFVEYANAGDSKAOKLLTGMPFGKVVATFYPLSAVKGYLQYTL 419
DB 361 LVPKENPGRGVFVEYANAGDSKAOKLLTGMPFGKVVATFYPLSAVKGYLQYTL 419

RESULT 4
AAO13515
ID AAO13515 PRELIMINARY; PRT; 419 AA.
AC AAO13515;
DT 02-MAR-2004 (Tremblrel. 27, Created)
DT 02-MAR-2004 (Tremblrel. 27, Last sequence update)

```

DT 02-MAR-2004 (TreeMrel. 27, Last annotation update)
DE KIS kinase.
GN KIS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCB1_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=12956;
RA Crook M.F., Boehm M., Nabel E.G.;
RL Submitted (Oct-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY180177; AAC13515.1; -
KW Kinase.
SQ SEQUENCE 419 AA; 46489 MW; 3BD3C06A9B2540 CRC64;

Query Match 98.9%; Score 2182; DB 2; Length 419;
Best Local Similarity 98.8%; Pred. No. 2,3e-157;
Matches 414; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAGSGCAGAEPRPLAEFGRLMQVOSRLGSSGSSAYVRYRCCNGSPGALROFLPPG 60
DB 1 MAGSGCAGAEPRPLAEFGRLMQVOSRLGSSGSSAYVRYRCCNGSPGALROFLPPG 60
QY 61 TTGAASAAEYGRKRAALEQOQHNVITLYGVFTTHSPVPSRCILLDLVSVSE 120
DB 61 TTGAASAAEYGRKRAALEQOQHNVITLYGVFTTHSPVPSRCILLDLVSVSE 120
QY 61 TTGAASAAEYGRKRAALEQOQHNVITLYGVFTTHSPVPSRCILLDLVSVSE 120
DB 61 TTGAASAAEYGRKRAALEQOQHNVITLYGVFTTHSPVPSRCILLDLVSVSE 120
QY 121 LILYSSHGCGSMNIOHCARDVLEALFLHGEVYADLKPRITWASNECFKLDPL 180
DB 121 LILYSSHGCGSMNIOHCARDVLEALFLHGEVYADLKPRITWASNECFKLDPL 180
QY 121 LILYSSHGCGSMNIOHCARDVLEALFLHGEVYADLKPRITWASNECFKLDPL 180
DB 121 LILYSSHGCGSMNIOHCARDVLEALFLHGEVYADLKPRITWASNECFKLDPL 180
QY 181 SFKEGNQDVYIOTDGYRAPEALQNCIAQAGLSQDTECTSAVDLSLGIILLEMFSGMK 240
DB 181 SFKEGNQDVYIOTDGYRAPEALQNCIAQAGLSQDTECTSAVDLSLGIILLEMFSGMK 240
QY 241 LKHTVRSQEWKANSIADHIFASKAVVNAIPAYHLRLIKSLHDDPSRRIPAEALC 300
DB 241 LKHTVRSQEWKANSIADHIFASKAVVNAIPAYHLRLIKSLHDDPSRRIPAEALC 300
QY 301 SPFFSIPFAHIDLVLPVRLNVLDLLEDEDEVEDVEDEKCECKIGVSVSL 360
DB 301 SPFFSIPFAHIDLVLPVRLNVLDLLEDEDEVEDVEDEKCECKIGVSVSL 360
QY 361 LVPKENGPGQVVEYANAGDSKAAQKLTGRMFDGKVVATFPPLSAVKGILYOTLL 419
DB 361 LVPKENGPGQVVEYANAGDSKAAQKLTGRMFDGKVVATFPPLSAVKGILYOTLL 419
DB 361 LVPKENGPGQVVEYANAGDSKAAQKLTGRMFDGKVVATFPPLSAVKGILYOTLL 419
DB 361 LVPKENGPGQVVEYANAGDSKAAQKLTGRMFDGKVVATFPPLSAVKGILYOTLL 419

RESULT 5
KIST MOUSE STANDARD; PRT; 419 AA.
ID KIST_MOUSE
AC P97133; O61775; O9CYT1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Serine/threonine-protein kinase Kist (EC 2.7.1.37) (Kinase interacting
DE with streptin).
GN Name=Kist; Synonyms=KIS;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCB1_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=57BL/6J; TISSUE=Embryo;
RC MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RX Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RX Nikaido I., Oasato N., Saito R., Yamanaoka I., Kiyosawa H.,
RX Nagai K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T.,
RX Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RX Schirml L.M., Kanapin A., Matsuda H., Batalov S., Batzel K.W.,
RX Blake J.A., Brand D., Brusic V., Chochia C., Corbani L.E., Cousins S.,

```

```

RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Kongaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglocz D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Nunata K., Okido T., Pavan W.J., Petrea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wymshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arawaka T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
RN (2)
RP SEQUENCE OF 1-414 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97435279; PubMed=9287318;
RA Maucuer A., Ozon S., Manceau V., Gavet O., Lawler S., Curmi P.,
RA Sobel A.;
RT "KIS is a protein kinase with an RNA recognition motif."
RL J. Biol. Chem. 272:23151-23156 (1997).
RN (3)
RP REVISIONS TO 52 AND 68-69.
RA Maucuer A.;
RL Submitted (Jul-2002) to the EMBL/GenBank/DBJ databases.
RN (4)
RP SEQUENCE OF 122-321 FROM N.A., AND INTERACTION WITH STATMIN.
RC TISSUE=Embryo;
RX MEDLINE=95241452; PubMed=7724523;
RA Maucuer A., Camonis J.H., Sobel A.;
RT "Statmin interaction with a putative kinase and coiled-coil-forming
RT protein domains."
RL Proc. Natl. Acad. Sci. U.S.A. 92:3100-3104 (1995).
CC -1- FUNCTION: Possesses kinase activity. May be involved in
CC trafficking and/or processing of RNA (By similarity).
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphorylated
CC -1- SUBUNIT: Interacts with streptin.
CC -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y10725; CA01714.2; -
DR EMBL: AK013347; BAB28802.1; -
DR EMBL: X62320; CAA57763.1; -
DR PIR: I48615; I48615.
DR HSSP: P26368; 100P.
DR MGD: MGI:1341908; Kist.
DR GO: GO:0005737; Cytoplasm; ISS.
DR GO: GO:0005634; C:Nucleus; ISS.
DR GO: GO:0005524; F:ATP binding; ISS.
DR GO: GO:0005515; F:protein binding; IPI.
DR GO: GO:0004674; F:protein serine/threonine kinase activity; ISS.
DR GO: GO:0003723; F:RNA binding; ISS.
DR GO: GO:0018105; P:peptidyl-serine phosphorylation; ISS.
DR InterPro: IPR011009; Kinase_like.

```

DR InterPro: IPR000719; Prot kinase.
 DR InterPro: IPR000504; RNA_rec_mot.
 DR InterPro: IPR008271; Ser_thr_pkin_AS.
 DR Pfam: PF00069; Kinase; 1.
 DR Pfam: PF00076; RM_1; 1.
 DR ProDom: PD000001; Prot_kinase; 1.
 DR SMART: SM00360; RM; 1.
 DR PROSITE: PS00107; PROTEIN KINASE ATP; FALSE_NEG.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
 DR PROSITE: PS50102; RM; 1.
 DR PROSITE: PS00030; RM_RNP_1; FALSE_NEG.
 DR ATP-binding; Nuclear protein; RNA-binding;
 KM Serine/chreonine-protein kinase; Transferase.
 FT DOMAIN 23 303
 FT NE_BIND 323 405 RNA-binding (RM).
 FT BINDING 29 37 ATP (By similarity).
 FT ACT_SITE 54 54 ATP (By similarity).
 FT CONFLICT 158 158 Proton acceptor (By similarity).
 FT CONFLICT 170 170 N -> D (in Ref. 4).
 FT CONFLICT 231 231 T -> I (in Ref. 3 and 4).
 SQ SEQUENCE 419 AA; 46477 MW; A91B307BA50F29A1 CRC64;

Query Match 98.7%; Score 2177; DB 1; Length 419;
 Best Local Similarity 98.6%; Pred. No. 5.5e-157;
 Matches 413; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MASSGCAMGAPRRFLFAFRRLNQVGRSLGSSSASYRRCCGNGSPFGARQRLPRG 60
 1 MASSGCAMGAPRRFLFAFRRLNQVGRSLGSSSASYRRCCGNGSPFGARQRLPRG 60

DB 61 TTGAASAAEYGRKERAALEQGHNIYTLGVFTIHSPPVPSRCLLELDVSVSE 120
 61 TTGAASAAEYGRKERAALEQGHNIYTLGVFTIHSPPVPSRCLLELDVSVSE 120

QY 121 LLLYSSHQGSMWMIQHCARDVLEALFLHREGVHADLPRNIIILSAENECFKLIDFGL 180
 121 LLLYSSHQGSMWMIQHCARDVLEALFLHREGVHADLPRNIIILSAENECFKLIDFGL 180

DB 121 LLLYSSHQGSMWMIQHCARDVLEALFLHREGVHADLPRNIIILSAENECFKLIDFGL 180
 121 LLLYSSHQGSMWMIQHCARDVLEALFLHREGVHADLPRNIIILSAENECFKLIDFGL 180

QY 181 SFPEGNQDVYKITTQDGRAPAEIQLNLAQGLQSDTECTSAVDLSGLIILMFSGSK 240
 181 SFPEGNQDVYKITTQDGRAPAEIQLNLAQGLQSDTECTSAVDLSGLIILMFSGSK 240

DB 181 SFPEGNQDVYKITTQDGRAPAEIQLNLAQGLQSDTECTSAVDLSGLIILMFSGSK 240
 181 SFPEGNQDVYKITTQDGRAPAEIQLNLAQGLQSDTECTSAVDLSGLIILMFSGSK 240

QY 241 LKHTVRSQEWKANSATIDHIFASKAVVNAIPAYHLRDLIKSLMDHDDPSRRIIPAEALC 300
 241 LKHTVRSQEWKANSATIDHIFASKAVVNAIPAYHLRDLIKSLMDHDDPSRRIIPAEALC 300

DB 241 LKHTVRSQEWKANSATIDHIFASKAVVNAIPAYHLRDLIKSLMDHDDPSRRIIPAEALC 300
 241 LKHTVRSQEWKANSATIDHIFASKAVVNAIPAYHLRDLIKSLMDHDDPSRRIIPAEALC 300

QY 301 SPFFSIPFAPIEDLVMLPTPVRLNLNVLDVLYGNEEYEDVDEYKECQKYGPIVSL 360
 301 SPFFSIPFAPIEDLVMLPTPVRLNLNVLDVLYGNEEYEDVDEYKECQKYGPIVSL 360

DB 301 SPFFSIPFAPIEDLVMLPTPVRLNLNVLDVLYGNEEYEDVDEYKECQKYGPIVSL 360
 301 SPFFSIPFAPIEDLVMLPTPVRLNLNVLDVLYGNEEYEDVDEYKECQKYGPIVSL 360

QY 361 LVKKNPGRQVVEYVANAAGDSKAAQKLTGRNPDGKVVATYPIISAYKRGYLYQTL 419
 361 LVKKNPGRQVVEYVANAAGDSKAAQKLTGRNPDGKVVATYPIISAYKRGYLYQTL 419

DB 361 LVKKNPGRQVVEYVANAAGDSKAAQKLTGRNPDGKVVATYPIISAYKRGYLYQTL 419
 361 LVKKNPGRQVVEYVANAAGDSKAAQKLTGRNPDGKVVATYPIISAYKRGYLYQTL 419

RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=99279253; PubMed=10349636;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44(1999).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=21085660; PubMed=11217851;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RA The FANTOM Consortium;
 RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN (4)
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=20499374; PubMed=11042159;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtration of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN (5)
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=20530913; PubMed=11076861;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Saeki N., Carninci P., Kono H., Akiyama Y., Nishi K., Katsunai T., Tashiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Wachihi M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN (6)
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P., Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W., Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T., Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kanakawa T., Katoh H., Kawai Y., Kojima Y., Kondo S., Kono H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M., Nishi K., Nomura K., Numazaki R., Ohno M., Ohse N., Okazaki Y., Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H., Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M., Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T., Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AK030152; BAC26812.1; -.
 DR HSSP: P26368; 100P.
 DR MGD: MGI:1341908; Kist.
 DR GO: GO:0005524; F-ATP binding; IEA.
 DR GO: GO:0016740; F-transportase activity; IEA.
 DR GO: GO:0004468; P-transportase activity; IEA.
 DR GO: GO:0004468; P-transportase activity; IEA.
 DR InterPro: IPR000719; Prot kinase.
 DR InterPro: IPR000504; RNA_rec_mot.
 DR Pfam: PF00069; Kinase; 2.
 DR Pfam: PF00076; RM_1; 1.
 DR ProDom: PD000001; Prot_kinase; 1.
 DR SMART: SM00360; RM; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.

DR	PROSITE: PS50102; RRM; 1.
KW	NAME: RRM
SC	SEQUENCE 330 AA; 37201 MW; 4164C0758BE7CB4C CRC64;
QY	Query Match 77.6%; Score 1711; DB 2; Length 330; Best Local Similarity 98.8%; Pred. No. 1.1e-121; Matches 326; Conservative 2; Mismatches 2; Indels 0; Gaps 0
DY	90 VTLGVFPIHSPNPNRCLLELLDVSVELLLSSHQGSGMMWIOHCARDVLEALFL 149 :::
DY	1 MTLGVFIHSPNPNRCLLELLDVSVELLLSSHQGSGMMWIOHCARDVLEALFL 60 :::
DY	150 HHEGVHADLKPRLIWSAENECEFKLIDFGISFKENGDVKKYIQTDGRAPAEALQNCIA 209 HHEGVHADLKPRLIWSAENECEFKLIDFGISFKENGDVKKYIQTDGRAPAEALQNCIA 120
DY	210 QAGLQSDPTECHSAVDLWSLGLITLLEMSGKMLKHTVRSQEWKANSALIIDHIFAKRAYVN 269 QAGLQSDPTECHSAVDLWSLGLITLLEMSGKMLKHTVRSQEWKANSALIIDHIFAKRAYVN 180
DY	270 AAIPIAYHLRDLIKSMHDDPSRRIPAEAMALCSFFSIFPAPIEDLVMLPTVRLRLNLVL 329 AAIPIAYHLRDLIKSMHDDPSRRIPAEAMALCSFFSIFPAPIEDLVMLPTVRLRLNLVL 240
DY	330 DDVYLGNEEYEDVDVEYBECQKGPVSLVLPENPGRGVFTVEYANAGDSKAAQKLL 389 DDVYLGNEEYEDVDVEYBECQKGPVSLVLPENPGRGVFTVEYANAGDSKAAQKLL 300
DY	390 TGRMPDGFVVAATFYPFLSAVYKRGVYKQTL 419 TGRMPDGFVVAATFYPFLSAVYKRGVYKQTL 330

Query	Best local similarity	11.5%	Score 254;	DB 2;	Length 505;
Matches	87;	Conservative	47;	Mismatches 129;	Indels 82;
					Gaps 11;
Qy	23	WQVSRLLGSSGSSASVAVRRCGNGSPGALRQFLPPTGTGAASAAEYGRKERAALEQ	82		
Db	22	FOVTEKELGHGAGYIVCAKYATGTGTGDAVATKVTNIFSKNILLCKRA----	135		
Qy	83	LOGHRNIVTLXGVTFTHFSPVPRCILLLELDV-SVSEILLYSGHSCSMMW----	135		
Db	78	FRGHKNTTCYLD-----MDIVDTNNRHELTLYEELMECDHGIIRSGQ	121		
Qy	136	-----OHCARDVLEALFLHHEGYVHADLKPRNILMSAENECEKLLIDFGLSFKEGNDV	189		
Db	122	LTPDAHYSFVYQILAGKYIHSADVLRDLKRGVLNVADCE-LKICDFGLARGSSDDE	180		
Qy	190	K-----YIQTDDYRPAPEALONCLAOAGLQSTECTSAVDLSGLIILEMFSKMLK	242		
Db	181	KNAGFLTEYVATRYRPAPEIML-----SFQSYTKAIDIMSVGCIILLELLGG--K	227		
Qy	243	HTVAGSQMKANSSALIDHI-----FASKRVNAAIPAYH--	276		
Db	228	PLFEGKNVVDQNLNLIHYLTGTPSEETLKRGSPRAQEVYRGLPMPKTPSTLPTANPE	287		
Qy	277	LRDLIKSMLEHDDPSERRIPAEWALCSPPFSIIPFAPIHEDLVMLPTP	321		
Db	288	ALDLEKMLAFDPARERTVEALHPLVLIKTIWHD--RDEPVCPTP	330		
RESULT 8					
ID	08WOL7	PRELIMINARY;	PRT;	882	AA.
AC	08WOL7	022155;			
DT	01-MAR-2002	(TREMblrel. 20, Created)			
DT	01-MAR-2002	(TREMblrel. 20, Last sequence update)			
DT	01-OCT-2004	(TREMblrel. 28, Last annotation update)			
DE	Minibrain kinase (Hypochemical protein T04C10.1).				
GN	Name=mbk-1; Synonym=T04C10.1;				
OS	Caenorhabditis elegans.				
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;				
OC	Rhabditidae; Pelodidae; Caenorhabditis.				
OX	NCBI_TaxID=6239;				
RN	1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=22505392; Pubmed=12618396;				
RA	Ratcl W.B., Mooman C., Laceyfield C.O., Lehrer J., Bartech D.,				
RA	Plasterk R.H., Kandel E.R., Hobert O.;				
RT	"Characterization of Caenorhabditis elegans homologs of the Down				
RL	syndrome candidate gene DYRK1A.";				
RN	Genetics 163:571-580(2003).				
RP	2]				
RX	SEQUENCE FROM N.A.				
RA	STRAIN=Briscot N2;				
RX	MEDLINE=99069613; Pubmed=9851916;				
RA	none;				
RT	"Genome sequence of the nematode C.elegans: A platform for				
RL	investigating biology.";				
RL	Science 282:2012-2018(1998).				
RP	3]				
RX	SEQUENCE FROM N.A.				
RA	STRAIN=Briscot N2;				
RA	Burton J.;				
CC	Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.				
CC	-I- SMILLARY: Belongs to the Ser/Thr protein kinase family.				
DR	EMBL; AY064464; AAL40874.1; -				
DR	EMBL; Z69885; CA93756.2; -				
DR	HSSP; P24941; IH0V.				
DR	Interact; 08WOL7; -				
DR	GO; GO:0005524; F:ATP binding; IEA.				
DR	GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.				
DR	GO; GO:0016740; F:transferase activity; IEA.				
DR	GO; GO:0006468; P:transferase amino acid phosphorylation; IEA.				

R	A	Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schler G.D.,
R	A	Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bat N.K.,
R	A	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
R	A	Diachenko L., Marutina K., Farmer A.A., Rubin G.M., Hong L.,
R	A	Stapleton M., Soares M.B., Bonaldi M.P., Casavant T.L., Scheetz T.E.,
R	A	Brownstein M.J., Usdin T.B., Toshguyji S., Carinci P., Prange C.,
R	A	Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
R	A	Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
R	A	Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.U., Huliy S.W.,
R	A	Villalón D.K., Wuzny D.M., Sodergren E.U., Lu X., Gibbs R.A.,
R	A	Fahy J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
R	A	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
R	A	Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
R	A	Rodriguez R.C., Grimwood J., Schultz J., Myers R.M., Butterfield Y.S.,
R	A	Kryzanski M.I., Skalska U., Smalius D.E., Scherch A., Schein J.E.,
R	A	Jones S.J., Maria M.A.,
R	T	"Generation and initial analysis of more than 15,000 full-length human
R	L	and mouse cDNA sequences.";
R	L	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
R	N	[2]
R	P	SEQUENCE FROM N.A.
R	P	TISSUE=Kidney;
R	C	MEDLINE=22341132; PubMed=12454917;
R	X	Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
R	A	Richardson P.;
R	T	"Genetic and genomic tools for Xenopus research. The NIH Xenopus
R	T	initiative.";
R	L	Dev. Dyn. 225:384-391(2002).
R	N	[3]
R	P	SEQUENCE FROM N.A.
R	P	TISSUE=Kidney;
R	A	Klein S., Strausberg R.,
R	L	Submitted (JUN-2004) to the EMBL/genbank/DBD databases.
R	C	-1 SIMILARITY: Belongs to the Ser/Thr protein kinase family.
D	R	EMBL; BC074183; AAB74183.1; -
D	R	InterPro; IPR011009; Kinase_lke.
D	R	InterPro; IPR002290; Ser_Thr_kinase.
D	R	InterPro; IPR008271; Ser_thr_pkin_AS.
D	R	InterPro; IPR01245; Tyr_kinase.
D	R	Pfam; PF00069; Pkinase; 1.
D	R	ProDom; PD000001; Prot_Kinase; 1.
D	R	SMART; SMO0220; S_TKC; 1.
D	R	SMART; SMO0219; TyKc; 1.
D	R	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
D	R	PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
D	R	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
D	R	ATP-binding; Hypothetical protein; Kinase;
D	R	Serine/threonine-protein kinase; Transferrase.
D	R	SEQUENCE 395 Aa; 43896 MW; 4C78BBBF828A2CC4 CRC64;
S	O	
Q	y	Query Match 10.9%; Score 240.5; DB 2; Length 395;
Q	y	Best Local Similarity 27.4%; Pred. No. 7.5e-10;
M	a	Matches .90; Conservative 48; Mismatches 108; Indels 83; Gaps 16
Q	y	4 SGCAWGAPRRFLAEFGRLMVOYQRSLGSGSSASV-----YYRCCGNFSPGPCA 52
D	b	9 SGACMKKK-----VEDIKRMFOFKETLGTGAFFSEVLAIEKETGKLFAVKCI-----PKKA 59
Q	y	53 LRQLPLPTTGAAASAAYEGPFKEKRAALEOLOGHNITLYLVGFTHFRSPNPSRCLL- 111
D	b	60 LK-----GKSSSIENEIAVLRIKTK-HENIVALEDIY-----ESPSHLYIWM 99
Q	y	112 -----ELLDVSVSELIIYSHQGCSMMMIQHCAADVLEALAFHHEGVHADLKPRNIL 165
D	b	100 QLVGGGLFD-RIVAKPGYTEKDASTL-----IRQVDAYSYLRLGIVRHDKLPKNLL 152
Q	y	166 WSAENECFKLI-DFGISFKEGNQDYKIYQ--TDGYRAPAEALQNCLAQAIGIQSTECTS 221
D	b	153 YFSQVEESKIMISDGLSKMEKGDMVSTAQTGPYVAPEV-----LQA-----KPYSK 201
Q	y	222 AVDLMSLGIIILEMWSGMKTKRTTASQEKMANSSAIIDHIPASKAVVNAAIPAY-----H 276

Db	202	ANDCMSIGIAVAILLCGY-----PPFYDENOSRLFEQLT--KADYEFDSPYMDISSES	255
Qy	277	LRDLIKSMHDDPSRRIPAEALCSPPS	305
Db	253	ADDFIQNIEMKDPNRRKYTCEQALRHPWIA	281
RESULT 11.			
Q9CA22	PRELIMINARY	PRT;	720 AA.
AC	Q9CA22		
DT	01-JUN-2001 (TREMBLrel. 17, Created)		
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)		
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)		
DE	Hypothetical protein T32E8.5.		
GN	Name=132E8.5;		
OC	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;		
OC	euroside II; Brassicales; Brassicaceae; Arabidopsis.		
OX	NCBI_TaxID=3702;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,		
RA	Maiti R., Rongning C.M., Koo H., Fujii C.Y., Uteback T.R.,		
RA	Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;		
RL	Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RA	Town C.D., Kaul S.;		
RL	Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL: AC012193; MAG51619.1; --		
DR	PIR: A96807; A96807.		
DR	GO: GO:0005524; F:ATP binding; IEA.		
DR	GO: GO:0004674; F:protein serine/threonine kinase activity; IEA.		
DR	GO: GO:0016740; F:transferase activity; IEA.		
DR	GO: GO:0006688; P:protein amino acid phosphorylation; IEA.		
DR	InterPro: IPR011009; Kinase like.		
DR	InterPro: IPR000719; Prot. Kinase.		
DR	InterPro: IPR002290; Ser. thr. kinase.		
DR	InterPro: IPR008271; Ser. thr. pkin_AS.		
DR	Pfam: PF00069; Pkinase; 1.		
DR	Prodom: PD000001; Prot. kinase; 1.		
DR	SMART: SM00220; S_TKc_1.		
DR	PROSITE: PS00107; PROTEIN KINASE ATP; UNKNOWN_1.		
DR	PROSITE: PS50011; PROTEIN KINASE DOM; 1.		
DR	PROSITE: PS01008; PROTEIN KINASE ST; UNKNOWN_1.		
KM	ATP-binding; Hypothetical protein; Kinase; Transferase.		
SO	SEQUENCE 720 AA; 80027 MW; F26A9A0B1C6D91DE CRC64;		
Query Match 10.9%; Score 240.5; DB 2; Length 720;			
Best local similarity 25.7%; Pred. No. 1.6e-09;			
Matches 105; Conservative 57; Mismatches 146; Indels 101; Gaps 17;			
Qy	11	EPFRPLEAFGRIMVQSHLSSGSSASVYRRCGNSPSPGALRQLPGPTGAASAAE	70
Db	331	DDPLFFKVGKYLQSLGIGSSSEVHKVY--SSDCTYALKKIKLKGRDYAT----	383
Qy	71	YGFKEKRALAEQLQGRHRIVLYGVFTIHSPVPSKCLLELDIVSE-----	120
Db	384	YGFCEIGIKLKIKKKNIIID-----IDYETTDKTLLEVLNQTMSNKGKRVKDSF	436
Qy	121	---LLLYS---SHQGSWM-----MICHQARDVLEALFLHREGYHADIK	160
Db	437	IYVWLEYGIDLAHMLSQKMEIEGSDRTIDENMLRPMYQQLIQAVNTIHEEIVHSDIK	496
Qy	161	PNRIILMSAENEFKILIDGSLFKSGNDVKYIQD-----GRAPAEALQNCIAAGIQ	214
Db	497	PANFL--IVRGLFKLIDGIA-KAINDDTNIRDSQVGLTSTWSPEAPMCNSEDENG--	551
Qy	215	SPTEGCSAVDLMSGIIILEMFSGMKLKHIVRSGEWKANSATIID--HIFASKAVYNAAI	272
Db	552	NITKCGRPDSILASCLIIYQVNYG-KRPAPADYKTF--AKKVTIDPNHEITYNQLSNP--	607

```

QY 273 PAYHLNDLIKSMHDDPSRRIPAEMLCSPFFSIPAPH----- 311
DB 608 ---WLDLWKCKCLAWBRNGWRRIPELLQHFLAPL-PHEPOYKTIKFLSLAESGSD 663
QY 312 -----IEDVLMPTPVLLRLNVLDDYLGNEEBYEDVEDVEEC 351
DB 664 DKANSWISOLEQLSNPAPLPR-----NDVLSRDQNOQLLSRVSLELC 706

RESULT 12
Q25758 PRELIMINARY; PRT; 765 AA.
ID Q25758;
AC Q25758;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, last annotation update)
DE Mitogen-activated protein kinase-related protein.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxId=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Dd2;
RA Lin D.T., Goldman N.D., Syin C.;
RT "Stage-specific expression of a Plasmodium falciparum protein related to the eukaryotic mitogen-activated protein kinases.";
RL Mol. Biochem. Parasitol. 78:0-0(10).
(2)
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Dd2;
RA Lin D.T., Goldman N.D., Syin C.;
RT Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
RL -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC EMBL: U36377; AAC47170.1; -.
DR HSSP: Q16539; IKV1.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0006688; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR011009; Kinase_like.
DR InterPro: IPR003527; MAP_kin.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR008271; Ser_thr_pkin_AS.
DR Pfam: PF00069; Pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKC; 1.
DR PROSITE: PS01351; MAPK; UNKNOWN_1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
KW SEQUENCE 765 AA; 90231 MW; 5BEC8EA6A4B0232 CRC64;

Query Match 10.9%; Score 240.5; DB 2; Length 765;
Best Local Similarity 25.1%; Pred. No. 1.7e-09;
Matches 103; Conservative 57; Mismatches 145; Indels 105; Gaps 19;
QY 23 WQVOSRLAGSSASVYVRCCGPGSP-----GALROFLPPTGGAASAEYGRKXR 77
DB 23 YDLKKVAGAGVAVKGRCKKNKIVAVKIGARQN-----CTDAQTRR-EL 71
QY 78 AALEQLQGRNIVTLTGVFTIHSFVPSRCLLLELDVSVSL-----LLYSHQCSMW 133
DB 72 IFLYELNGHNDIITKMDVIR-----AKDNDIYLLFDMEETDLEHVIKADLLEETIK----- 123
QY 134 MIOHCARDVLEALFLHHEGVADLKPRVILMSAENEC-FKLIDRGLS-----FKEN 186
DB 124 --KTIIVQLRAKTIHSGGLRDIKPSNITL--VNSECHIKVADGLASISTHYNENK 179
QY 187 ODV--KYIOTDGYRABEALQNCIAQAGLSQDTECTSAVDLMSLGIITLLEMPGML----- 241

```

```

DB 180 VPIITDYATRWYRAPEI-----LLGSTHYTEDVDWMSLGLCINGELLGKPLFTG 229
QY 242 KATVRSOE-----WKAANSATIID--HIFASAVVN-AAIPAYHNR-----DLI 281
DB 230 NSTMNQLEKIIVIGVGNPKKDIEDIRSPAEKTISSFPDILKKAKIDCIYASNESL 289
QY 282 KSMHDDPSRRIPAEML-----CSPFFSIPAPHIEDVLMPTPVLLRLNVLDDYLGNE 337
DB 290 EKLQFNPSKRIISANALMKHYVEEFHSIIDPFCRHIIITP----- 331

QY 338 EYEDVEDVEKECCQKPGVSVLVPKENPGAGQFVEYANAGDSKAAOK 387
DB 332 -----INDTKYRVNFYRNVYFVIMRRNKFHSNVL-----VQGESKKEK 372

RESULT 13
Q84VX4 PRELIMINARY; PRT; 777 AA.
ID Q84VX4;
AC Q84VX4;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, last annotation update)
DE Atlg77720.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eupside II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Hsuang V.W.,
RA Ishida J., Jones T., Kamliya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G., Yuan S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BR004636; AA042882.1; -.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0006688; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR011009; Kinase_like.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR008271; Ser_thr_pkin_AS.
DR Pfam: PF00069; Pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE: PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
KW ATP-binding; Kinase; Transferase.
KW SEQUENCE 777 AA; 86376 MW; F2C25DA9609A878F CRC64;

Query Match 10.9%; Score 240.5; DB 2; Length 777;
Best Local Similarity 25.7%; Pred. No. 1.7e-09;
Matches 105; Conservative 57; Mismatches 146; Indels 101; Gaps 17;
QY 11 EPPRFLEAFGRMVOVSRGSSASVYVRCCGPGSPGALROFLPPTGGAASAAE 70
DB 388 DDDLFFKNGVKIYQALGKIGSGSSSEVKNVT---SSDCIYALKIKYKAGRYATN----- 440
QY 71 YGFRKRAALEQLQGRNIVTLTGVFTIHSFVPSRCLLLELDVSVSE----- 120
DB 441 YGFCOEIGYIKLKKKTNITQL-----IDYEVTDKTLDEVLNGTWSNKKDGRVKEGDF 493
QY 121 ---LLYS-----SHQCSMW-----MIOHCARDVLEALFLHHEGVADLK 160
DB 494 IYVLELGEIDLAHMLSQKWRBIEGSDRTIDENMWLFYWOQLQAVNTHIERIRVSDLK 553
QY 161 PNILMSAENECFKLIDRGLSPKEGNQDVKYIOTD-----GYRABEALQNCIAQAGLQ 214

```

```

Db      554 PANFL--LVRGFLKLDIFGIA--KAINSDTTNIORDSQVGLTSYSPFAWCMESDENG-- 608
Qy      215 SDTECSAVDWLMSIGIILEMFSQMKLKHVRSQEWKANSALIID--HIFASAVVNAAI 272
Db      609 NPTKCRPSGIVSLGGLLYOMVYG--RTPPADYKTFW--AKFKVITDENHETITYNQLSNP-- 664
Qy      273 PAYHLADLTKSMHDDPSRRIPAEMLACSPFSPISPPAPH----- 311
Db      665 ---WILDIMKKCLAWPDRQWRIRPELQHPFLAPPI--PHEPOVKTIKIFSLIASEGSD 720
Qy      312 -----IEDLVMLPTPVRLINLVDDYLCNESEYEDVEDVKEEC 351
Db      721 DKANSMISQLEQLSNPAPLPR-----NDVIDRSDONQOLLSRVSELC 763

RESULT 14
ID      094656 PRELIMINARY; PRT; 826 AA.
AC      094656;
DT      01-FEB-1997 (TEMBLrel. 02, Created)
DT      01-FEB-1997 (TEMBLrel. 02, Last sequence update)
DE      01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE      Mitogen-activated protein kinase 1, serine/threonine protein
       kinase.
DE      Name=pfmap1;
OS      Plasmodium falciparum.
OC      Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX      NCBI_TaxID=5833;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=IT04;
RA      Doering C.;
RL      Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC      -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR      EMBL; X82646; CNA57972.1; -.
DR      PIR; JCS153; JCS153.
DR      HSSP; Q16539; 1KV1.
DR      GO; GO:000524; F:ATP binding; IEA.
DR      GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR      GO; GO:0016740; F:transferase activity; IEA.
DR      GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR      InterPro; IPR011009; Kinase like.
DR      InterPro; IPR003527; MAP kin.
DR      InterPro; IPR000719; Prot_kinase.
DR      InterPro; IPR002290; Ser_thr_kinase.
DR      pfam; PF00069; Pkinase; 1.
DR      SMART; SM00220; S_TKc; 1.
DR      PROSITE; PS01351; MARK; UNKNOWN 1.
DR      PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR      PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR      PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW      ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ      SEQUENCE 826 AA; 96645 MW; A66BEF7E734A38A CRC64;

Query Match 10.9%; Score 240.5; DB 2; Length 826;
Best Local Similarity 25.1%; Pred. No. 1.9e-09;
Matches 103; Conservative 57; Mismatches 145; Indels 105; Gaps 19;

Qy      23 WQVOSRLSGSSAVYRVRCGNGSPSP-----GALRQFLPPTGTGAASAAYGFRKER 77
Db      23 YDLIKKVGKAGVGVFRGRCKKKNIVAVKIRGATQN-----CTDAQRTFR--EL 71

```

```

Qy      78 AALBOLQGRNIVLTGYVFTTHSPNVPSRCILLELVVSSEL-----LIYSHGCGSMW 133
Db      72 IFLYEINGDNIETIKMDVIK--AKNDNDIYLIFPFMETDHEVIKADLEIRK----- 123
Qy      134 MIOHCARDVLEALAFHHEGVHADLKPRNITLSAENEC--FLDIPFGS-----FKEGN 186
Db      124 --KVIITVQLRALRKIIHSGGLIHRDIKPSNII--VNSECHIKVADFGIARISITHVENK 179
Qy      187 QDV--KYIOTDGYRAPEALONCLAQIQSPTCSTAVDWLSIGIILEMFSQMKL--- 241
Db      180 VPILDYVATRWYRAPEI-----LIGSTHYTEVDVMSIGCINGELCGKPLFTG 229
Qy      242 KATVRSQ-----WKANSALIID--HIFASAVVN-AAIPAYHR-----DLI 281
Db      230 NSTMNQLEKTIQVICKPKPKDIEDIRSPFAEKTISSFVDLKKKKNLCYRASNESLDL 289
Qy      282 KSMHDDPSRRIPAEML-----CSPFSPISPPAPHIEDLVMLPTPVRLINLVDDYLCNE 337
Db      290 EKLQFPNSKRISAEMLKHKVEEFSHIIDEPFCRHIIITP----- 331
Qy      338 EEYEDVEDVKEECQYGPVVSILVKNPGRQVFEVYANAGSKAAQK 387
Db      332 -----INNTKXRVNFRNVVYFVIMRRNKKFHSNVL-----NGGSKKEK 372

RESULT 15
ID      081LF0 PRELIMINARY; PRT; 914 AA.
AC      081LF0;
DT      01-MAR-2003 (TEMBLrel. 23, Created)
DT      01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DE      01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE      Mitogen-activated protein kinase 1.
GN      ORFNames=PF14_0294;
OS      Plasmodium falciparum (isolate 3D7).
OC      Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX      NCBI_TaxID=36329;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      MEDLINE=22255705; Pubmed=12368864;
RA      Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
       Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
       Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kys S.,
       Chan M.S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
       Perlea M., Allen J., Selengut J., Hait D., Mather M.W., Valdivia A.B.,
       Martin D.M., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
       McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
       Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
       Frazer C.M., Barrett B.;
RL      "Genome sequence of the human malaria parasite Plasmodium
       falciparum."
RC      Nature 419:498-511(2002).
CC      -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR      EMBL; AE014820; AAN36907.1; -.
DR      HSSP; Q16539; 1KV1.
DR      GO; GO:000524; F:ATP binding; IEA.
DR      GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR      GO; GO:0016740; F:transferase activity; IEA.
DR      GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR      InterPro; IPR011009; Kinase like.
DR      InterPro; IPR003527; MAP kin.
DR      InterPro; IPR000719; Prot_kinase.
DR      InterPro; IPR002290; Ser_thr_kinase.
DR      InterPro; IPR008271; Ser_thr_pkin_AS.
DR      pfam; PF00069; Pkinase; 1.
DR      SMART; SM00220; S_TKc; 1.
DR      PROSITE; PS01351; MARK; UNKNOWN 1.
DR      PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR      PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR      PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW      ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ      SEQUENCE 914 AA; 107277 MW; CA4F486AB051F472 CRC64;

```

Query Match	10.9%	Score 240.5;	DB 2;	Length 914;
Best Local Similarity	25.1%;	Pred. No. 2.1e-09;		
Matches 103; Conservative	57;	Mismatches 145;	Indels 105;	Gaps 19;

```

QY      23  QOVQSRILSGSSASAYRVACCGNPGSP-----GALRGFLPBTGTAASAAAYGRKER 77
Db      23  YDILKKVKGAGVVPFKGCKKKNKIVAAKKI FGAQN-----CTDAQRTR-El 71
QY      78  AALBEOLOGRHNIIVTVLYGVFTIHPSPNVPSPRCLELLELDVSVEL-----LTSYSHQCSNM 133
Db      72  IFLYELANGHDNIILKMDVIK-----AKNDNDIYILFDMEPMDIHEVIRKADLLEELHK----- 123
QY      134  MIQHCAADVLELALFLHEEGYHADLKPNILMSAENEC-FKLDIFGLS-----FREGN 186
Db      124  --KYLIIYOLLRAKXIHSGGLIHRDIPKSNIL--VNSECHIKVADGELGARSISTHVENK 179
QY      187  QDV--KYIQTDSYRAPEALQWNCLOAQAGSDPECSAVDWLSGIILLEMFGMKL--- 241
Db      180  VPILDYATIRWYRAPEI-----LIGSTHYIEDVDWMSLGGCIMGELLCGRPLFTG 229
QY      242  KHTVRSOE-----WKANSSAIID--HIEASKAVVN-AAIPAYHLR-----DLI 281
Db      230  NSTNNQLEKLIQVIGKPNKKDIEDIRSPAEEKIISFVLDKKNNLIDICYKANSESIDL 289
QY      282  KSMHDDPSRRIPAEAL-----CSPEFSIPPAAPHIEDLVMLPTPVIRLLANVLDDYLGNE 337
Db      290  EKTLQFNESKRISAENALGKRYVEEHSIIDEPTCRHIIITP----- 331
QY      338  BEYEDVVEVEKECQKYPVSLVLYKPKNGRQGVVEYVANGSDSKAAOK 387
Db      332  -----INDTKTRVFNTRNVVIFVILRRNKFTSNV-----NQGESKKEEK 372

```

Search completed: November 29, 2004, 11:46:53
Job time : 123.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2004, 11:26:47 / Search time 109.5 Seconds
(without alignments)
1372.672 Million cell updates/sec

Title: US-10-798-532-4
Perfect score: 2206
Sequence: 1 MAGSCGAMGAEPPRFLEAFG.....VATFPLSAVKRGVLYQTLL 419

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues
Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp1980a:*
2: geneseqp1990a:*
3: geneseqp2000a:*
4: geneseqp2001a:*
5: geneseqp2002a:*
6: geneseqp2003a:*
7: geneseqp2003bs:*
8: geneseqp2004a:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	ID	Description
1	2195	99.5	419 4 AAG65767	Aag65767 Human pro
2	2195	99.5	419 5 ABP64797	Abp64797 Human pro
3	2190.5	99.3	418 3 AAY70306	Aay70306 Human tra
4	2188	99.2	419 4 AAU03534	Aau03534 Human pro
5	2188	99.2	419 4 AAE11769	Aae11769 Human kin
6	2187.5	99.2	418 3 AAY70305	Aay70305 Human wil
7	2183	99.0	419 7 ADD48093	Add48093 Rat Prote
8	483.5	21.9	137 5 ABP64662	Abp64662 Human ORF
9	240.5	10.9	638 3 AAG38502	Aag38502 Arabidops
10	240.5	10.9	720 3 AAG38501	Aag38501 Arabidops
11	240.5	10.9	777 3 AAG38500	Aag38500 Arabidops
12	236.5	10.7	544 2 AAW85028	Aaw85028 CDK2-gree
13	235	10.7	298 7 ADD21398	Add21398 Golden ha
14	234.5	10.6	433 8 AAB19094	Aab19094 Protein e
15	234.5	10.6	433 8 ADP87472	Adp87472 S cerevis
16	234	10.6	376 3 AAG04672	Aag04672 Arabidops
17	233	10.6	435 6 AAE31916	Aae31916 Yeast Hog
18	232.5	10.5	367 3 AAB21137	Aab21137 Rat stres
19	232	10.5	435 6 AAE31917	Aae31917 Yeast Hog
20	232	10.5	435 6 AAE31922	Aae31922 Yeast Hog
21	232	10.5	435 6 AAE31921	Aae31921 Yeast Hog
22	232	10.5	435 6 AAE31914	Aae31914 Yeast Hog
23	232	10.5	435 6 AAE31913	Aae31913 Yeast Hog
24	232	10.5	544 2 AAW85029	Aaw85029 Green flu
25	232	10.5	568 2 AAW75791	Aaw75791 Human pro

26	232	10.5	568 3 AAB00059	Aab00059 hYAK3-2 p
27	232	10.5	568 4 AAE02013	Aae02013 Human YAK
28	232	10.5	587 7 AAB57639	Ab57639 Human DYK
29	232	10.5	588 2 AAW75790	Aaw75790 Human pro
30	232	10.5	588 3 AAY68792	Aay68792 Amino aci
31	232	10.5	588 3 AAB00058	Aab00058 hYAK3-2 a
32	232	10.5	588 4 AAE02012	Aae02012 Human YAK
33	232	10.5	801 7 ADD93285	Add93285 p53-Yak3
34	231.5	10.5	371 3 AAG45420	Aag45420 Arabidops
35	231.5	10.5	384 2 AAR89411	Aar89411 Stress ac
36	231.5	10.5	426 2 AAR89410	Aar89410 Stress ac
37	231.5	10.5	426 5 AAB85009	Ab85009 Pain regu
38	231.5	10.5	549 8 ADI45327	Adi45327 Rice isop
39	231	10.5	298 2 AAW95689	Aaw95689 Homo sap1
40	231	10.5	298 7 ADD21402	Add21402 Chinese h
41	231	10.5	298 7 ADD21406	Add21406 Rat cycl1
42	231	10.5	298 7 ADD21394	Add21394 Mouse cyc
43	231	10.5	416 2 AAW06795	Aaw06795 High Osmo
44	231	10.5	435 6 AAE31910	Aae31910 Yeast Hog
45	231	10.5	435 6 AAE31915	Aae31915 Yeast Hog

ALIGNMENTS

RESULT 1
ID AAG65767 standard; protein; 419 AA.
XX
AC AAG65767;
XX
DT 07-JAN-2002 (first entry)
XX
DE Human protein kinase polypeptide 13887.
XX
KW Protein kinase; 3714; 16742; 23546; 13887; cancer; bone disorder; human; cytosolic; antiinflammatory; immunosuppressive; cardiac; hepatotrophic; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200173050-A2.
XX
PD 04-OCT-2001.
XX
PF 23-MAR-2001; 2001WO-US009483.
XX
PR 24-MAR-2000; 2000US-0191846P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Meyers R;
XX
XX WPI; 2001-611632/70.
DR N-PSDB; AAI66828, AAI66829.
XX
PT New human protein kinase polypeptides; 3714, 16742, 23546 and 13887, useful in diagnosis of cancer or cellular proliferation or to screen for polypeptide modulators useful to treat such conditions.
PT
PT
XX
PS Claim 4; Fig 10A-C; 16pp; English.
XX
XX The invention provides novel human protein kinase polypeptides; 3714, 16742, 23546 and 13887 and nucleic acid molecules encoding them. The protein kinase polypeptides can be expressed by standard recombinant methodology; 3714, 16742, 23546 or 13887 nucleic acids and polypeptides are useful for diagnostic and screening methods to identify subjects (at risk of) having cancer or cellular proliferation and/or differentiation disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and modulators are useful for the treatment of cancer, particularly colon cancer or cellular proliferation and/or differentiation disorders. Other disorders associated with 3714, 16742, 23546 or 13887 expression or

CC activity that can be treated include bone related disorders, inflammatory
 CC disorders, autoimmune diseases, cardiovascular disorders and liver
 CC diseases. The present sequence represents a human protein kinase
 CC polypeptide 13867

XX Sequence 419 AA;

Query Match 99.5%; Score 2195; DB 4; Length 419;
 Best Local Similarity 99.5%; Pred. No. 5.3e-229;
 Matches 417; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAGSGCAGWAEPPRFLFAFGRIMQVQSRIGSSSASVYRRCGNGSPGALRQFLPPG 60
 DB 1 MAGSGCAGWAEPPRFLFAFGRIMQVQSRIGSSSASVYRRCGNGSPGALRQFLPPG 60
 QY 61 TTGAASAAEYGRKERALAEQOGHNTVLTGVTTHFSVNVPSRCLLELDVSVSE 120
 DB 61 TTGAASAAEYGRKERALAEQOGHNTVLTGVTTHFSVNVPSRCLLELDVSVSE 120
 QY 121 LLLYSSHQGCSMMWIOHCARDVLEALAFHHEGYVHADLPRNIIWSAENECFKLIDFGL 180
 DB 121 LLLYSSHQGCSMMWIOHCARDVLEALAFHHEGYVHADLPRNIIWSAENECFKLIDFGL 180
 QY 181 SFKEGQDVYKTIQTDGRAPAEALQNCIAQAGIQTDECTSAVDLWSLGIILLEMPSGK 240
 DB 181 SFKEGQDVYKTIQTDGRAPAEALQNCIAQAGIQTDECTSAVDLWSLGIILLEMPSGK 240
 QY 241 LKHTVRSQEMKANSATIDHIFASKAVVNAIPAYHLRDLIKSMLHDDPSRRIPAEALC 300
 DB 241 LKHTVRSQEMKANSATIDHIFASKAVVNAIPAYHLRDLIKSMLHDDPSRRIPAEALC 300
 QY 301 SPFFSIPFAPHIEDLVMLPTPVRLNLNVDDYLGNEEYEDVDEVDKECQKYGVS 360
 DB 301 SPFFSIPFAPHIEDLVMLPTPVRLNLNVDDYLGNEEYEDVDEVDKECQKYGVS 360
 QY 361 LVPEKMPGRGVFVEYANAGDSKAAQKLLTGMPDQGFVATFYPLSAYKRGYLTLL 419
 DB 361 LVPEKMPGRGVFVEYANAGDSKAAQKLLTGMPDQGFVATFYPLSAYKRGYLTLL 419

RESULT 2
 ABP64797
 ID ABP64797 standard; protein; 419 AA.

XX ABP64797;

AC 25-FEB-2003 (first entry)

DE Human protein SBQ ID 457.

XX Human; expressed sequence tag; EST; haematopoietic disorder;
 KW central nervous system disease; viral infection;
 KW peripheral nervous system disease; non-healing wound; infectious disease;
 KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;
 KW fungal infection; autoimmune disorder; coagulation disorder; noctropic;
 KW antiallergic; antiinflammatory; immunosuppressive; neutroprotective;
 KW cytostatic; haemostatic; virucide; antibacterial; fungicide;
 KW immunostimulant; cerebroprotective.

XX Homo sapiens.

XX WO200259260-A2.

PD 01-AUG-2002.

XX 16-NOV-2001; 2001MO-US042950.

PR 17-NOV-2000; 2000US-00714936.

PA (HYSB-) HYSB INC.

PI Tang YT, Goodrich RM, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
 Ren F, Xue AJ, Yang Y, Wehman T, Drmanac RT;

XX WPI; 2002-590824/63.
 DR N-PSDB; ABQ99383.
 XX New isolated polynucleotide, useful in research, diagnostic or
 PT therapeutic methods, e.g. preventing or treating disorders involving
 PT aberrant protein expression or biological activity.

PS Claim 20; SEQ ID NO 457; 394bp; English.

XX The present invention relates to novel human coding sequences (ABQ99268-
 CC ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in
 CC therapeutic, diagnostic, and research methods. The polynucleotides may be
 CC used in the field of molecular biology as hybridization probes, primers
 CC for PCR, for chromosome and gene mapping, for the recombinant production
 CC of protein, or in generation of anti-sense DNA or RNA. The
 CC polynucleotides are useful in diagnostics as expressed sequence tags
 CC (ESTs) for identifying expressed genes or for physical mapping of the
 CC human genome. The proteins may be used as molecular weight markers, or as
 CC nutritional sources or supplements. The proteins may be used to maintain
 CC and expand cell population in a totipotent or pluripotent state
 CC useful for re-engineering damaged or diseased tissues, transplantation,
 CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The
 CC polynucleotides and proteins are useful for preventing, treating or
 CC ameliorating disorders involving aberrant protein expression or
 CC biological activity, e.g. haematopoietic disorders, central/peripheral
 CC nervous system diseases, mechanical and traumatic disorders, non-healing
 CC wounds, immune deficiencies and disorders, infectious diseases caused by
 CC viral, bacterial or fungal infection, autoimmune disorders, allergic
 CC reactions and conditions, coagulation disorders, or cancer. The
 CC polynucleotide sequences of the invention were assembled from ESTs
 CC isolated mainly by sequencing by hybridization, and in some cases,
 CC sequences obtained from one or more public databases. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPRO at
 CC ftp.wipro.int/pub/published_pct_sequences

XX Sequence 419 AA;

Query Match 99.5%; Score 2195; DB 5; Length 419;
 Best Local Similarity 99.5%; Pred. No. 5.3e-229;
 Matches 417; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAGSGCAGWAEPPRFLFAFGRIMQVQSRIGSSSASVYRRCGNGSPGALRQFLPPG 60
 DB 1 MAGSGCAGWAEPPRFLFAFGRIMQVQSRIGSSSASVYRRCGNGSPGALRQFLPPG 60
 QY 61 TTGAASAAEYGRKERALAEQOGHNTVLTGVTTHFSVNVPSRCLLELDVSVSE 120
 DB 61 TTGAASAAEYGRKERALAEQOGHNTVLTGVTTHFSVNVPSRCLLELDVSVSE 120
 QY 121 LLLYSSHQGCSMMWIOHCARDVLEALAFHHEGYVHADLPRNIIWSAENECFKLIDFGL 180
 DB 121 LLLYSSHQGCSMMWIOHCARDVLEALAFHHEGYVHADLPRNIIWSAENECFKLIDFGL 180
 QY 181 SFKEGQDVYKTIQTDGRAPAEALQNCIAQAGIQTDECTSAVDLWSLGIILLEMPSGK 240
 DB 181 SFKEGQDVYKTIQTDGRAPAEALQNCIAQAGIQTDECTSAVDLWSLGIILLEMPSGK 240
 QY 241 LKHTVRSQEMKANSATIDHIFASKAVVNAIPAYHLRDLIKSMLHDDPSRRIPAEALC 300
 DB 241 LKHTVRSQEMKANSATIDHIFASKAVVNAIPAYHLRDLIKSMLHDDPSRRIPAEALC 300
 QY 301 SPFFSIPFAPHIEDLVMLPTPVRLNLNVDDYLGNEEYEDVDEVDKECQKYGVS 360
 DB 301 SPFFSIPFAPHIEDLVMLPTPVRLNLNVDDYLGNEEYEDVDEVDKECQKYGVS 360
 QY 361 LVPEKMPGRGVFVEYANAGDSKAAQKLLTGMPDQGFVATFYPLSAYKRGYLTLL 419
 DB 361 LVPEKMPGRGVFVEYANAGDSKAAQKLLTGMPDQGFVATFYPLSAYKRGYLTLL 419

RESULT 3

AAU0306
 ID AAY70306 standard; protein; 418 AA.
 AC AAY70306;
 DT 06-JUN-2000 (first entry)
 XX Human transdominant mutant serine/threonine kinase KIS (hKIS) protein.
 DE
 XX KIS; human; hKIS; serine/threonine kinase; transdominant mutant; p27;
 KW G1 phase; cell proliferation; cyclin-dependent kinase inhibitor; CKI;
 KW modulator; treatment; cell proliferative disease; vascular disorder;
 KW gene therapy; atherosclerosis; restenosis.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 54 /note= "Wild type Lys substituted with Arg"
 FT Misc-difference 166 /note= "Encoded by AATCAG"
 FT Misc-difference 186 /note= "Encoded by AATCAG"
 XX
 PN WO20001165-A1.
 XX
 PD 02-MAR-2000.
 XX
 PF 20-AUG-1999; 99WO-US018903.
 XX
 PR 21-AUG-1998; 98US-0097710P.
 XX
 PA (NABE/) NABEL G J.
 PA (NABE/) NABEL E G.
 XX
 PI Nabel GJ, Nabel EG;
 XX
 XX WPI; 2000-237648/20.
 DR N-PSDB; AAZ51356.
 XX
 XX Novel serine/threonine kinase hKIS polynucleotides and polypeptides used
 PT for inhibiting the cyclin kinase inhibitor p27, and so alter cell
 PT proliferation.
 XX
 XX Example 1; Page 59-61; 70pp; English.
 PS
 XX The present protein sequence is the transdominant mutant human KIS
 CC (hKIS), constructed by site directed mutagenesis. A single nucleotide
 CC substitution (A to G) results in a lysine to arginine change in the
 CC protein sequence. hKIS is a serine/threonine kinase, that acts as an
 CC inhibitory kinase of cyclin-dependent kinase inhibitor (CKI), p27. hKIS
 CC controls cell proliferation and is localised predominantly in the
 CC nucleus. It binds to CKI p27 and inhibits its ability to arrest cells in
 CC G1 phase. The hKIS sequences are used to modulate cell proliferation and
 CC treat cell proliferative and vascular diseases. The polynucleotide
 CC sequence may be used in gene therapy to treat vascular disorders such as
 CC restenosis or atherosclerosis
 CC
 XX
 SQ Sequence 418 AA;
 Query Match 99.3%; Score 2190.5; DB 3; Length 418;
 Best Local Similarity 99.8%; Pred. No. 1.6e-228;
 Matches 418; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1 MAGSGCMAEPRPRLEAFGRMLQVQSRGLGSSASVYRRCGCCNGSPGALRQRLPPG 60
 DB 1 MASSGGAMGAEPRPRLEAFGRMLQVQSRGLGSSASVYRRCGCCNGSPGALRQRLPPG 60
 QY 61 TTGAASAAEYGRKERRALEQOGHRNITVLKGVFTIHSPVNPVSRCLLELDVSVSE 120
 DB 61 TTGAASAAEYGRKERRALEQOGHRNITVLKGVFTIHSPVNPVSRCLLELDVSVSE 120
 QY 121 LLLYSSHQGGSMWMIQHCADVLEALAFHHEGYVHADLKPRNIIWASNECEFKLIDFGL 180
 DB 121 LLLYSSHQGGSMWMIQHCADVLEALAFHHEGYVHADLKPRNIIWASNECEFKLIDFGL 180

QY 181 SPEKGNQDVYKTIQTDGYRAPEALONCLAQAGSDTECTSAVDLSIGIILLEMFGMK 240
 DB 181 SPEKGN-DVKYIQTQGYRAPEALONCLAQAGSDTECTSAVDLSIGIILLEMFGMK 239
 QY 241 LKHTVRSQEWKANSNAIIDHIFASKAVVNAIPAYHLRDLIKSLHDDPSRRIPAEMLC 300
 DB 240 LKHTVRSQEWKANSNAIIDHIFASKAVVNAIPAYHLRDLIKSLHDDPSRRIPAEMLC 299
 QY 301 SPFFSIPFAPHIEDLVMLPTPLRLNLNTDDYLCNEEYEVVDVYEECKYGPVSL 360
 DB 300 SPFFSIPFAPHIEDLVMLPTPLRLNLNTDDYLCNEEYEVVDVYEECKYGPVSL 359
 QY 361 LVPKENPGRGQVFEVYANAGDSKAOKLLTGMFPDGKPVATFYPASAKRGYLYOTLL 419
 DB 360 LVPKENPGRGQVFEVYANAGDSKAOKLLTGMFPDGKPVATFYPASAKRGYLYOTLL 418
 RESULT 4
 AAU03534
 ID AAU03534 standard; protein; 419 AA.
 XX
 AC AAU03534;
 XX
 DT 12-SEP-2001 (first entry)
 XX
 DE Human protein kinase #34.
 XX
 XX Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
 KW metabolic disorder; immune related disease; neurological disorder;
 KW neurodegenerative disorder; inflammatory disorder; infectious disease;
 KW reproductive disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200138503-A2.
 XX
 PD 31-MAY-2001.
 XX
 PF 22-NOV-2000; 2000MO-US032085.
 XX
 PR 24-NOV-1999; 99US-0167482P.
 XX
 PA (SUGR-) SUGEN INC.
 XX
 PI Plowman GD, Whyte D, Manning G, Sudarshanam S, Martinez R;
 PI Flanagan P, Clary D;
 XX
 DR WPI; 2001-343950/36.
 XX
 DR N-PSDB; AAS06734.
 XX
 XX Nucleic acids encoding human kinase polypeptides, useful for preventing
 PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
 PT neuronal-associated diseases, and microbial infections.
 XX
 PS Claim 7; Fig 2; 433pp; English.
 XX
 XX AAU03501-AAU03557 represent novel human protein kinases #1-57. The novel
 CC protein kinases have been identified as members of the tyrosine or
 CC serine/threonine kinase (PTK and STK) families. The polynucleotides
 CC encoding protein kinases and the polypeptides may be used in the
 CC prevention, diagnosis and treatment of diseases associated with
 CC inappropriate kinase expression. For example, they may be used to treat
 CC cancers (especially cancers of haematopoietic origin), cardiovascular
 CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
 CC immune related diseases (e.g. rheumatoid arthritis), neurological
 CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
 CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
 CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
 CC Additionally, polynucleotides encoding protein kinases may be used for
 CC gene therapy and as DNA probes in diagnostic assays. The protein kinase
 CC polypeptides may be used as antigens in the production of antibodies
 CC against the protein kinases and in assays to identify modulators of

CC protein kinase expression and activity
 XX Sequence 419 AA:
 Query Match 99.2%; Score 2188; DB 4; Length 419;
 Best Local Similarity 99.3%; Pred. No. 3.1e-228;
 Matches 416; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAGSGCAGWAEPPRFLFAFGRLWQVSRGSSSASVYVRCCNGPSPGALROFLPPG 60
 DB 1 MAGSGCAGWAEPPRFLFAFGRLWQVSRGSSSASVYVRCCNGPSPGALROFLPPG 60
 QY 61 TTGAASAAYGFRKRAALBOLQGHRIIVTLYGVTTFHSPVNSRCLLELDVSVSE 120
 DB 61 TTGAASAAYGFRKRAALBOLQGHRIIVTLYGVTTFHSPVNSRCLLELDVSVSE 120
 QY 121 LLLYSHOGSGMMWICHCADVLEALAFHHEGYVHADLPKRIILMSAENECKLIDFGL 180
 DB 121 LLLYSHOGSGMMWICHCADVLEALAFHHEGYVHADLPKRIILMSAENECKLIDFGL 180
 QY 181 SFKEGNQDVYIOTDGYRAPEALONCLAQGLQSDPECTSAVDLMSLGIILLEMESGK 240
 DB 181 SFKEGNQDVYIOTDGYRAPEALONCLAQGLQSDPECTSAVDLMSLGIILLEMESGK 240
 QY 241 LKHTVRSQEWKANSIAIIDIHFASKAVVNAIPAYHLRDLIKSLHDDPSRRIPAEALC 300
 DB 241 LKHTVRSQEWKANSIAIIDIHFASKAVVNAIPAYHLRDLIKSLHDDPSRRIPAEALC 300
 QY 301 SPFFSIPPAHIEDLWMLPTPVRLNLVDDLYLENEEYEDVEDKECKYGPVSL 360
 DB 301 SPFFSIPPAHIEDLWMLPTPVRLNLVDDLYLENEEYEDVEDKECKYGPVSL 360
 QY 361 LVPKPMRGQVVEVEYANAGDSKAQKLTGRMFDGKVVATFYPISAYKRGYOTLL 419
 DB 361 LVPKPMRGQVVEVEYANAGDSKAQKLTGRMFDGKVVATFYPISAYKRGYOTLL 419

RESULT 5
 AAE11769
 ID AAE11769 standard; protein; 419 AA.
 XX
 AC AAE11769;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human kinase (PKIN) -3 protein.
 XX
 KW Human kinase; PKIN; gene therapy; adenocarcinoma; immune disorder; gout;
 KW cancer; allergy; sarcoma; leukaemia; acquired immune deficiency syndrome;
 KW AIDS; Addison's disease; microbial infection; inflammation; osteoporosis;
 KW atherosclerosis; cardiovascular disease; myocardial infarction; anaemia;
 KW myasthenia gravis; cirrhosis; cataract; growth and development disorder;
 KW seizure disorder; pulmonary embolism; Gaucher's disease; lipid disorder;
 KW lipid storage disease; Pick's disease; Tay-Sachs disease; renal disease;
 KW asthma; obesity; restorative therapy; cytostatic; immunomodulatory;
 KW antimicrobial; cardiovascular; antiinflammatory; vaccine.
 XX
 OS Homo sapiens.
 XX
 PH Key
 XX
 FT Domain
 FT Location/Qualifiers
 FT 23..304
 FT /note="Eukaryotic protein kinase domain"
 FT 48..289
 FT /note="Protein kinase domain"
 FT 74..293
 FT /note="Protein kinase domain"
 FT 108..241
 FT /note="Protein kinase domain"
 FT 148..166
 FT /note="Tyrosine kinase catalytic domain"
 FT 223..245
 FT /note="Tyrosine kinase catalytic domain"
 FT 223..241

FT /label=Transmembrane_domain
 FT Domain 273..295
 FT /note="Tyrosine kinase catalytic domain"
 FT Region 305..344
 FT /note="Serine/threonine protein kinase"
 FT Region 319..404
 FT /note="Ribonucleoprotein repeat"
 FT Region 320..412
 FT /note="Splicing factor like protein"
 PN MO200181555-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001MO-US012992.
 XX
 PR 20-APR-2000; 2000US-0199021P.
 PR 28-APR-2000; 2000US-0200226P.
 PR 05-MAY-2000; 2000US-0202339P.
 PR 11-MAY-2000; 2000US-0203505P.
 PR 18-MAY-2000; 2000US-0205644P.
 PR 26-MAY-2000; 2000US-0207739P.
 PR 01-JUN-2000; 2000US-0208795P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Yue H, Gandhi AR, Tribouley CM, Kearney L, Griffln JA, Nguyen DB;
 PI Bandman O, Lu DM, Lal P, Burford N, Khan FA, Walla NK, Yao MG;
 PI Patterson C, Burrill JD, Marcus GA, Zingler KA, Reippon SA, Lu Y;
 PI Policky JL, Thornton M, Tang YT, Hafalia A, Elliott VS, Baughn MR;
 PI Walsh RT, Ramkumar J, Borowsky ML, Au-Young J, Hillman JL,
 PI Gururajan R;
 XX
 DR MPI; 2001-611740/70.
 DR N-PSDB; AAD18818.
 XX
 PT Human kinases and nucleic acids, useful for preventing diagnosing and
 PT treating cancers, inflammation and immune disorders.
 XX
 PS Claim 1; Page 125-126; 166pp; English.
 XX
 CC The present invention relates to human kinases (PKIN) and the nucleic
 CC acids encoding them. PKIN is used as vaccine and in gene therapy. PKIN is
 CC used in the prevention, diagnosis and treatment of diseases cancers,
 CC adenocarcinoma, leukaemia, sarcoma, immune disorder, Addison's disease,
 CC acquired immune deficiency syndrome (AIDS), anaemia, asthma, allergies,
 CC gout, microbial infections, cardiovascular disease and/or inflammation;
 CC myasthenia gravis, atherosclerosis, cirrhosis, osteoporosis, myocardial
 CC infarction, cataract, growth and development disorder, seizure disorder,
 CC pulmonary embolism, Gaucher's disease, lipid disorder, lipid storage
 CC disease, Pick's disease, Tay-Sachs disease, renal disease and obesity.
 CC PKIN may be used to treat disorders associated with decreased PKIN
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of PKIN by expressing inactive proteins or to
 CC supplement the patient's own production of PKIN. PKIN nucleic acids may be
 CC used to produce the PKIN polypeptide, by inserting the nucleic acids into
 CC a host cell and culturing the cell to express the protein. PKIN nucleic
 CC acid and its complementary sequences may also be used as DNA probes in
 CC diagnostic assays to detect and quantitate the presence of similar
 CC nucleic acid sequences in samples and therefore which patients may be in
 CC need of restorative therapy. The present sequence is human PKIN-3 protein
 CC
 XX
 SQ Sequence 419 AA;
 Query Match 99.2%; Score 2188; DB 4; Length 419;
 Best Local Similarity 99.3%; Pred. No. 3.1e-228;
 Matches 416; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAGSGCAGWAEPPRFLFAFGRLWQVSRGSSSASVYVRCCNGPSPGALROFLPPG 60
 DB 1 MAGSGCAGWAEPPRFLFAFGRLWQVSRGSSSASVYVRCCNGPSPGALROFLPPG 60
 QY 61 TTGAASAAYGFRKRAALBOLQGHRIIVTLYGVTTFHSPVNSRCLLELDVSVSE 120

```

Db      61 TTGAASAAEYGRKERAALEQLQGHNIIVTLGVFTIHSPVPSRCLLELDVSVSE 120
Qy      121 LLIYSSHOGCSMMWIOHCARDVLEALAFHHEGYVADLKPRNIIWSAENECFKLIDFGL 180
Db      121 LLIYSSHOGCSMMWIOHCARDVLEALAFHHEGYVADLKPRNIIWSAENECFKLIDFGL 180
Qy      181 SFKEGNQDVYKIQTDGYRAPEAELOHCAQAGLOSDTECTSAVDLMSLGIILLEMFSGMK 240
Db      181 SFKEGNQDVYKIQTDGYRAPEAELOHCAQAGLOSDTECTSAVDLMSLGIILLEMFSGMK 240
Qy      241 LKHTVRSQEWKANSALIDHIFASKAVVNAIPAYHLRDLIKSMHDDPSRRIPAEMLC 300
Db      241 LKHTVRSQEWKANSALIDHIFASKAVVNAIPAYHLRDLIKSMHDDPSRRIPAEMLC 300
Qy      301 SPFFSIPFAPIHEDLVMLPTPVRLNVLDDYLGNEEYEDVVEDEKCECKYGPVSL 360
Db      301 SPFFSIPFAPIHEDLVMLPTPVRLNVLDDYLGNEEYEDVVEDEKCECKYGPVSL 360
Qy      361 LVPKENPGRGOVFEVYANAGDSKAAOKLITGRMFDGKFPVATFPLSAVKRGYLYOTLL 419
Db      361 LVPKENPGRGOVFEVYANAGDSKAAOKLITGRMFDGKFPVATFPLSAVKRGYLYOTLL 419

```

RESULT 6
ID AAY70305 standard; protein; 418 AA.
XX AAY70305;

06-JUN-2000 (first entry)

Human wild type serine/threonine kinase KIS (hKIS) protein.

KIS; human; hKIS; serine/threonine kinase; cell proliferation; G1 phase;
cyclin-dependent kinase inhibitor; CKI; p27; modulator; treatment;
cell proliferative disease; vascular disorder; gene therapy; restenosis;
atherosclerosis.

OS Homo sapiens.

Key Location/Qualifiers
FT Misc-difference 186 /note="Encoded by AATCAG"

PN MO20001165-A1.

PD 02-MAR-2000.

PF 20-AUG-1999; 99MO-US018903.

PR 21-AUG-1998; 98US-0097710P.

PA (NABE/) NABEL G J.
PA (NABE/) NABEL E G.

PI Nabel GJ, Nabel EG;

XX WPI, 2000-237648/20.

DR N-PSDB; AAZ51355.

PT Novel serine/threonine kinase hKIS polynucleotides and polypeptides used
for inhibiting the cyclin kinase inhibitor p27, and so alter cell
proliferation.

XX Example 1; Page 57-58; 70pp; English.

XX The present protein sequence is the wild type human KIS (hKIS), isolated
from a human B-cell library, using a yeast two hybrid screening system.
XX hKIS is a serine/threonine kinase, that acts as an inhibitory kinase of
XX cyclin-dependent kinase inhibitor (CKI), p27. hKIS controls cell
XX proliferation and is localised predominantly in the nucleus. It binds to
XX CKI p27 and inhibits its ability to arrest cells in G1 phase. It has an

CC amino terminal serine/threonine kinase consensus region and a carboxy
CC terminus with 42% sequence similarity to huzAF5, a subunit of the
CC splicing factor U2AF. The hKIS sequences are used to modulate cell
CC proliferation and treat cell proliferative and vascular diseases. The DNA
CC sequence may be used in gene therapy to treat vascular disorders such as
CC restenosis or atherosclerosis

XX Sequence 418 AA;

Query Match 99.2%; Score 2187.5; DB 3; Length 418;
Best Local Similarity 99.5%; Pred. No. 3.5e-228;
Matches 417; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

```

Qy      1 MAGSGCAGMGAEPFRPEAFGRILMVOVSRLGSGSSAVYVRCCNGSPGALROPLPG 60
Db      1 MAGSGCAGMGAEPFRPEAFGRILMVOVSRLGSGSSAVYVRCCNGSPGALROPLPG 60
Qy      61 TTGAASAAEYGRKERAALEQLQGHNIIVTLGVFTIHSPVPSRCLLELDVSVSE 120
Db      61 TTGAASAAEYGRKERAALEQLQGHNIIVTLGVFTIHSPVPSRCLLELDVSVSE 120
Qy      121 LLIYSSHOGCSMMWIOHCARDVLEALAFHHEGYVADLKPRNIIWSAENECFKLIDFGL 180
Db      121 LLIYSSHOGCSMMWIOHCARDVLEALAFHHEGYVADLKPRNIIWSAENECFKLIDFGL 180
Qy      121 LLIYSSHOGCSMMWIOHCARDVLEALAFHHEGYVADLKPRNIIWSAENECFKLIDFGL 180
Db      121 LLIYSSHOGCSMMWIOHCARDVLEALAFHHEGYVADLKPRNIIWSAENECFKLIDFGL 180
Qy      181 SFKEGNQDVYKIQTDGYRAPEAELOHCAQAGLOSDTECTSAVDLMSLGIILLEMFSGMK 240
Db      181 SFKEGN-DVKYIOTDGYRAPEAELOHCAQAGLOSDTECTSAVDLMSLGIILLEMFSGMK 239
Qy      241 LKHTVRSQEWKANSALIDHIFASKAVVNAIPAYHLRDLIKSMHDDPSRRIPAEMLC 300
Db      240 LKHTVRSQEWKANSALIDHIFASKAVVNAIPAYHLRDLIKSMHDDPSRRIPAEMLC 299
Qy      301 SPFFSIPFAPIHEDLVMLPTPVRLNVLDDYLGNEEYEDVVEDEKCECKYGPVSL 360
Db      300 SPFFSIPFAPIHEDLVMLPTPVRLNVLDDYLGNEEYEDVVEDEKCECKYGPVSL 359
Qy      361 LVPKENPGRGOVFEVYANAGDSKAAOKLITGRMFDGKFPVATFPLSAVKRGYLYOTLL 419
Db      360 LVPKENPGRGOVFEVYANAGDSKAAOKLITGRMFDGKFPVATFPLSAVKRGYLYOTLL 418

```

RESULT 7

ID ADD48093 standard; protein; 419 AA.

AC ADD48093;

DT 29-JAN-2004 (first entry)

DE Rat Protein AAC53031, SEQ ID NO 13791.

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

OS Rattus norvegicus.

PN MO2003016475-A2.

PD 27-FEB-2003.

PF 14-AUG-2002; 2002MO-US025765.

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-033347P.

PA (GENO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

PI Woolf C, D'urso D, Befort K, Coetigan M;

XX WPI; 2003-268312/26.

DR GENBANK; AAC53031.

XX New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 XX

PS Claim 1; Page, 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (SNII), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPD at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 419 AA:

Query Match 99.0%; Score 2183; DB 7; Length 419;
 Best Local Similarity 98.8%; Pred. No. 1.le-227;
 Matches 414; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MASSGCAMGAEPRPFLFAFRRLWQVSRGSSSASVYRRCGCCGSPGALRQFLPPG 60
 DB 1 MASSGCAMGAEPRPFLFAFRRLWQVSRGSSSASVYRRCGCCGSPGALRQFLPPG 60
 QY TTGAASAAEYGFRRKRAALEQLQGHNRITLVYGVFTIHSPVPSRCLLELLDVSVSE 120
 DB TTGAASAAEYGFRRKRAALEQLQGHNRITLVYGVFTIHSPVPSRCLLELLDVSVSE 120
 QY 61 TTTAAASAAEYGFRRKRAALEQLQGHNRITLVYGVFTIHSPVPSRCLLELLDVSVSE 120
 DB 61 TTTAAASAAEYGFRRKRAALEQLQGHNRITLVYGVFTIHSPVPSRCLLELLDVSVSE 120
 QY 121 LLLYSSHQGCSMMWMIQHICARDVLEALAFLLHHEGVVHADLPRNITLWAENECFKLIDFGL 180
 DB 121 LLLYSSHQGCSMMWMIQHICARDVLEALAFLLHHEGVVHADLPRNITLWAENECFKLIDFGL 180
 QY 181 SFKEGQDVYVYITDGRARAEALQNLCAQAGOSDPECTSAVDLWSLGIILLEMSSGMK 240
 DB 181 SFKEGQDVYVYITDGRARAEALQNLCAQAGOSDPECTSAVDLWSLGIILLEMSSGMK 240
 QY 241 LKKTIVSOEWKASSAIIIDHIFASKAVVNAIPAYHLRDLIKMLHDDPSRRIPAEALC 300
 DB 241 LKKTIVSOEWKASSAIIIDHIFASKAVVNAIPAYHLRDLIKMLHDDPSRRIPAEALC 300
 QY 301 SPFFSIPFAPHIDLVMLPTPVRLRLNLVDDDYLGNEEYEDVYEDVKEECQKYGPVSL 360
 DB 301 SPFFSIPFAPHIDLVMLPTPVRLRLNLVDDDYLGNEEYEDVYEDVKEECQKYGPVSL 360
 QY 361 LVKXENPGRQVVEVYANAGDSKAQQLTGRMPDQGFVATYPLSAYKRGYLYOTLL 419
 DB 361 LVKXENPGRQVVEVYANAGDSKAQQLTGRMPDQGFVATYPLSAYKRGYLYOTLL 419

RESULT 8
 ABP64662
 ID ABP64662 standard; protein; 137 AA.

XX ABP64662;
 AC
 XX
 DT 04-NOV-2002 (first entry)
 XX
 DE Human ORF1032.

XX Cytostatic; Cardiac; Anti-allergic; Immunosuppressive; Vascular;
 KW Antiinflammatory; gene therapy; human; ORFX; atherogenic; platelet;
 KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
 KW cancer; cardiovascular disease; allergy; autoimmune disease;
 KW wound healing; blood coagulation disorder; inflammatory disorder.

XX Homo sapiens.

XX US2002082206-A1.

XX 27-JUN-2002.

XX 30-MAY-2001; 2001US-00867550.

XX 30-MAY-2000; 2000US-0208427P.

XX (LEACH/) LEACH M D.
 PA (MEHR/) MEHRABAN F.
 PA (CONLEY/) CONLEY P B.
 PA (TOPPER/) TOPPER J N.
 PA (LAW D.) LAW D.

PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;

XX WPI; 2002-626554/67.

DR N-PSDB; ABQ99225.

PT New polypeptide designated ORFX are present in human atherogenic cells
 PT and are useful to prevent and treat ORFX-associated disorders including
 PT cancer, allergy, wound healing or autoimmune, cardiovascular or
 PT inflammatory disease.

XX Claim 10; SEQ ID NO 2064; 78pp; English.

XX The present invention relates to novel human ORFX polypeptides and their
 CC coding sequences (ABP63631-ABP64661 and ABQ98194-ABQ99267). The sequences
 CC were discovered in human atherogenic cells, in particular in platelets
 CC and human umbilical vein endothelial cells (HUVEC) and are expressed in
 CC many other tissues as well. Atherogenic cells are cells which have the
 CC potential to develop atherosclerotic plaques. The ORFX polypeptides and
 CC nucleic acids are useful for treating or preventing a pathological
 CC condition associated with an ORFX-associated disorder, e.g. cancer,
 CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood
 CC coagulation disorders or inflammatory disorders. Note: The sequence data
 CC for this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from the USPTO web site at
 CC seqdata.uspto.gov/sequence.html?docID=2002082206

XX Sequence 137 AA:

Query Match 21.9%; Score 483.5; DB 5; Length 137;
 Best Local Similarity 80.0%; Pred. No. 1e-43;
 Matches 96; Conservative 1; Mismatches 12; Indels 11; Gaps 1;

QY 63 GAAASAAEYGFRRKRAALEQLQGHNRITLVYGVFTIHSPVPSRCLLELLDVSVSELL 122
 DB 29 GRGAALFSGGLRAAVA-----VTLGVFTIHSPVPSRCLLELLDVSVSELL 77
 QY 123 LYSHQGCSMMWMIQHICARDVLEALAFLLHHEGVVHADLPRNITLWAENECFKLIDFGLSF 182
 DB 78 LYSHQGCSMMWMIQHICARDVLEALAFLLHHEGVVHADLPRNITLWAENECFKLIDFGLSF 137

RESULT 9
 AAG38502
 ID AAG38502 standard; protein; 638 AA.

XX AC AAG38502;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 47508.
XX DE Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121625P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 04-MAY-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.

PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0143542P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143524P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145813P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148175P.
PR 10-AUG-1999; 99US-0148177P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148355P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.

PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160880P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 25-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 10.9%; Score 240.5; DB 3; Length 638;
Best Local Similarity 25.7%; Pred. No. 3e-16;
Matches 105; Conservative 57; Mismatches 146; Indels 101; Gaps 17;

QY 11 EPPRFLEAFGRMLNOVOSRLGSSASAYRYRCGNPSPGALROFLPPTGTAASAE 70
DB 249 DPDLFRKVNCKLYORLAKTSSGSSSEVHKYI---SSCTTYALKKTKLGRDYATA---- 301
QY 71 YGFRKRAALQLOQHNRITLVGVFTIHPSPNVPSRCLLELDIVSVSE----- 120
DB 302 YGFCQELIGYKKUKKKNITQL-----IDYEVTDKTLLEQVLTNGTMSVKDGRVEDGF 354
QY 121 ---LLIYS---SHQGSNM-----MICHCAADVLEALAFHHEGYHADLK 160
DB 355 IYVAVLEGEIDLAMLSQKWRREIEGSDRTIDENWLKRYWOOILQAVVTIHEERIVHSDLK 414
QY 161 PRRIILSAEHECKRLIDFGSLFKGNODVXYIOTD-----GYRAPAEHLONCLAOAGLQ 214
DB 415 PANFL--LVNGFLKLDIFGIA-KAINSDTTNIGORDSOVGLTSLMSPAFACNESDENG-- 469

QY 215 SPECTSAVDLWSLGIILLEMFSGMLKHTVRSQGEWKANSATID--HIPAKAVVNAAI 272
DB 470 NTIKGGRPSDIWSLACILYQWYVG-RTPFADYKTFW-AKFYITDPNHETYNQLSNP-- 525
QY 273 PAYHRLDIKSMHLNDPSPRRIPAEEMALCSPPFSIFPAPH----- 311
DB 526 ---WLDLWKCKCLADNRQWRIPELLQHPFLAPPI-PHEPOVKTIKLFSLIAESGSD 581
QY 312 -----IEDLVMLPFPEVLRLLNVLDLDYLGNEEYEDVEDYKEEC 351
DB 582 DKANSMISGLEQLLSNPAPLPR-----NDVDSRDNQNGQLSRSELC 624

RESULT 10
AAG38501
ID AAG38501 standard; protein; 720 AA.
AC AAG38501;
XX
XX 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 47507.
XX KM Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
XX PR 05-MAR-1999; 99US-0123180P.
XX PR 09-MAR-1999; 99US-0123548P.
XX PR 23-MAR-1999; 99US-0125788P.
XX PR 25-MAR-1999; 99US-0126264P.
XX PR 29-MAR-1999; 99US-0126785P.
XX PR 01-APR-1999; 99US-0127462P.
XX PR 06-APR-1999; 99US-0128234P.
XX PR 08-APR-1999; 99US-0128714P.
XX PR 16-APR-1999; 99US-0129845P.
XX PR 19-APR-1999; 99US-0130077P.
XX PR 21-APR-1999; 99US-0130449P.
XX PR 23-APR-1999; 99US-0130510P.
XX PR 28-APR-1999; 99US-0130891P.
XX PR 30-APR-1999; 99US-0131449P.
XX PR 30-APR-1999; 99US-0132048P.
XX PR 04-MAY-1999; 99US-0132407P.
XX PR 05-MAY-1999; 99US-0132484P.
XX PR 06-MAY-1999; 99US-0132486P.
XX PR 06-MAY-1999; 99US-0132487P.
XX PR 07-MAY-1999; 99US-0132863P.
XX PR 11-MAY-1999; 99US-0134256P.
XX PR 14-MAY-1999; 99US-0134218P.
XX PR 14-MAY-1999; 99US-0134219P.
XX PR 14-MAY-1999; 99US-0134219P.
XX PR 14-MAY-1999; 99US-0134219P.
XX PR 14-MAY-1999; 99US-0134570P.
XX PR 18-MAY-1999; 99US-0134768P.
XX PR 19-MAY-1999; 99US-0134941P.
XX PR 20-MAY-1999; 99US-0135124P.
XX PR 21-MAY-1999; 99US-0135353P.
XX PR 24-MAY-1999; 99US-0136229P.
XX PR 25-MAY-1999; 99US-0136021P.
XX PR 27-MAY-1999; 99US-0136392P.
XX PR 28-MAY-1999; 99US-0136782P.
XX PR 01-JUN-1999; 99US-0137222P.
XX PR 03-JUN-1999; 99US-0137528P.

PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138694P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140659P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142290P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.

PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149246P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151067P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155489P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157533P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 10.9%; Score 240.5; DB 3; length 720;
Best Local Similarity 25.7%; Pred. No. 3.7e-16;

Matches 105; Conservative 57; Mismatches 146; Indels 101; Gaps 17;
QY 11 EPPRFLEAFGRLLQVQSRLLSGSSASAYRYRCGCPSPGALROFLPPTGAAASAE 70
Db 331 DPDLFFKRNCKLQRLKGISSGSSSEVHKYI---SSDCTTYALKIKLKGRDYATA---- 393
QY 71 YGFRKRAALBOLQGHNRITLVGVFTIHPSNVPSPKLLLELDVSVSE----- 120
Db 384 YGFCOEIGYKLLKQKTNITQL-----IDVEYTDKTLLOEVLTNGTSMKDGKVEDGF 436
QY 121 ---LLLYS---SHGCSMW-----MICHGADVLEALAFHHEGVADLK 160
Db 437 IYVNLVEYGEIDLAMLSQKWRLEIGSDRRTIDENWLLRYWOQITLQAVNTIHEERIVHSDLK 496
QY 161 PRNIIWSAENECFKLIDFGLSFKGNQDVYKIOTD-----GYRAPEALQNCIAQGLQ 214
Db 497 PANFL--LVRGFKLIDFGIA-KAINSDTNIQDQGVGLTSMSPBAFMCNSDENG-- 551
QY 215 SDTECTSAVDLWLSGIILLEMFSQMKLKHTVRSQEWKANSAAIID--HIPASKAVVNAAI 272
Db 552 NTKICGRPSDIMSIGLCITLYQVYG-RTPFADYKTFW-AKFKVITDPNHEITYNQLSNP-- 607
QY 273 PAVHADLIYSMLHDDPSRRIRPAMALCSPFSTIPFAPH----- 311
Db 608 --WLDLMKCKCLAMPDRNQWRIPELLQHPFLAPPI-PHEPQYKTIKLFSLIAESCGSD 663
QY 312 -----IEDLVMLPTVLRLLNVLDLDYLQNEEYEDVVDVKEEC 351
Db 664 DRANSMISQLEQLSNAPLPFPR-----NDVLSRDQNOGLLRVSLC 706
RESULT 11
AAG38500
ID AAG38500 standard; protein; 777 AA.
AC AAG38500;
XX
XX 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 47506.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
XX
XX EPI033405-A2.
XX
XX
PD 06-SEP-2000.
XX
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.

PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134570P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135533P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139763P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144332P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.

PR 22-JUL-1999; 99US-0145192P.
 PR 23-JUL-1999; 99US-0145145P.
 PR 23-JUL-1999; 99US-0145218P.
 PR 23-JUL-1999; 99US-0145224P.
 PR 26-JUL-1999; 99US-0145276P.
 PR 27-JUL-1999; 99US-0145913P.
 PR 27-JUL-1999; 99US-0145918P.
 PR 27-JUL-1999; 99US-0145919P.
 PR 28-JUL-1999; 99US-0145951P.
 PR 02-AUG-1999; 99US-0146386P.
 PR 02-AUG-1999; 99US-0146388P.
 PR 03-AUG-1999; 99US-0146389P.
 PR 04-AUG-1999; 99US-0147038P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 05-AUG-1999; 99US-0147302P.
 PR 05-AUG-1999; 99US-0147192P.
 PR 06-AUG-1999; 99US-0147260P.
 PR 06-AUG-1999; 99US-0147303P.
 PR 06-AUG-1999; 99US-0147416P.
 PR 09-AUG-1999; 99US-0147493P.
 PR 09-AUG-1999; 99US-0147935P.
 PR 10-AUG-1999; 99US-0148171P.
 PR 11-AUG-1999; 99US-0148319P.
 PR 12-AUG-1999; 99US-0148341P.
 PR 13-AUG-1999; 99US-0148565P.
 PR 13-AUG-1999; 99US-0148684P.
 PR 16-AUG-1999; 99US-0149368P.
 PR 17-AUG-1999; 99US-0149175P.
 PR 18-AUG-1999; 99US-0149426P.
 PR 20-AUG-1999; 99US-0149722P.
 PR 20-AUG-1999; 99US-0149723P.
 PR 20-AUG-1999; 99US-0149929P.
 PR 23-AUG-1999; 99US-0149802P.
 PR 23-AUG-1999; 99US-0149930P.
 PR 25-AUG-1999; 99US-0150566P.
 PR 26-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 27-AUG-1999; 99US-0151080P.
 PR 30-AUG-1999; 99US-0151303P.
 PR 31-AUG-1999; 99US-015138P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 07-SEP-1999; 99US-0152363P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 24-SEP-1999; 99US-0155486P.
 PR 28-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158029P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.

PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161922P.
 PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 10.9%; Score 240.5; DB 3; Length 777;
 Best Local Similarity 25.7%; Pred. No. 4, 2e-16;
 Matches 105; Conservative 57; Mismatches 146; Indels 101; Gaps 17;

QY 11 EPPRFLEAAGRLMOYQSRGSGSSAVYVRRCGNGSPGALROFLPPTGAASAAE 70
 DB 388 DPDLFFKNGKLYQKLGKISGSGSEVHKV---SSDCTIYALKKIKLGRDYATA---- 440
 QY 71 YGFRKERALEQLQGRNIVTLYGVTIHFSPNPSRCLLELDVSVSE----- 120
 DB 441 YGFCQEIYGLKTKKKTNIQL-----IDYEVTDKTLQELVNGTMSNKGKVEDGF 493
 QY 121 ---LLIYS---SHQCSMW-----MIOCRADYELMALHHEGYHADLK 160
 DB 494 IYVLEYGEIDLAHMLQKMBREIEGSDRTIDENWLRFYWOQLQAVNIHERIYHSLK 553
 QY 161 PNIIISANBECFKLIDPGLSPFEGNOVKYIQTD-----GYRPEALQNCCLAQAGIQ 214
 DB 554 PANFL--LVRGFLKIDPGLA-KAINSDTNIQRDSQVGTLSYMPFAFMCHESENG-- 608
 QY 215 SDPECTSAVDLWLSGIIILEMPSGMLKRTVRSQEMKANSaIID--HIFASKAVVNAI 272
 DB 609 NTIKGGRPSDWSLQICILYQWVYG-RTPFADYKTFW-AKFVITDPNHETYNQSNP-- 664
 QY 273 PAVHLRDLIKSMHDDPSRRIRPAEMALCSPPFSIFPAH----- 311
 DB 665 ---WLIDLMKKCLAMDRCQWRIPRILQHPFLAPPI--PHEPOVKTIKLFSLIABSCGSD 720
 QY 312 -----IEDLVNLPVPVLRNLVNDVDDYGNBEEYEDVEDYKERC 351
 DB 721 DKANSMISQLEQLLSNPAPLPR-----NDVIDSRQNOQLSRVSELC 763

RESULT 12
 ID AAW85028 standard; protein; 544 AA.
 XX AAW85028;
 AC AAW85028;
 XX 17-OCT-2003 (revised)
 DT 08-FEB-1999 (first entry)
 XX CDK2-green fluorescent protein usion product.
 DE Human; CDK2 gene; fusion protein; green fluorescent protein; GFP;
 XX intracellular signalling; chimera.
 OS Aegnoorea victoria.
 OS Homo sapiens.
 OS Chimeric.
 XX MO9845704-A2.
 XX 15-OCT-1998.
 PD 07-APR-1998; 98WO-DK000145.
 PF 07-APR-1998; 98WO-DK000145.
 XX 07-APR-1997; 97DK-00000392.

XX (NOVO) NOVO-NORDISK AS.
 PA Thastrup O, Petersen Bjorn S, Tullin S, Kasper A, Scudder K;
 PI WPI, 1998-594491/50.
 XX N-PSDB; AA071073.
 DR
 XX Determining effect on signalling pathways in live cells from
 PT redistribution of lumophores - specifically fusions of green
 PT fluorescent protein with a signalling component, and new apparatus,
 PT particularly for identifying toxins and potential therapeutic agents.
 PS
 XX Example 12; Page 208-209; 326pp; English.
 CC The present sequence represents a human CDK2-green fluorescent protein
 CC (GFP) fusion protein. The fusion protein is used in an assay to exemplify
 CC the invention. The specification describes how quantitative information
 CC about the influence of a molecule on a cellular response is obtained by
 CC recording the variation, caused by the molecule, on mechanically intact
 CC living cells, in the spatially distributed light emitted from a
 CC lumophore present in the cells. The variation in light emission is
 CC processed to provide information that correlates spatial distribution to
 CC the degree of the molecule. The method is used to identify agents that
 CC (in)directly affect intracellular signalling, especially to screen for
 CC potential therapeutic agents or toxins, and to identify new drug targets.
 CC (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 544 AA;

Query Match 10.7%; Score 236.5; DB 2; Length 544;
 Best Local Similarity 26.7%; Pred. No. 6.5e-16;
 Matches 111; Conservative 53; Mismatches 154; Indels 97; Gaps 18;

QY 23 MVOGSLGSGSSASVYRVRCGNGSPGALRQF-LPFGTGAASAAYGPKERAALE 81
 DB 4 FQKVEKIGEGTGVVYKAK--NKLGEVVALKKIRLDTEBVPSTAL-----REISLTK 56
 QY 82 QLOGHNRITVLYGVFTIHSPNVPBSRCLLELDVSVSELLYSSHOGCSMMMIQHCARD 141
 DB 57 EL-NHNPNIYKLDLV--IHTEKTL---YLVEFLHQDLKKFMDSALVGTIPLPIKSYLQ 110
 QY 142 VLEALFLHHEGVYADLKPRNIIWSAENECFKLIDFGLSFKGNQDVKY---IQTDGVR 198
 DB 111 LLOGLAFLCHSHRVLRDLKPNQNLINTEG-AIKLADFGIARAFGVPRVYTHEVVTLMYR 169
 QY 199 APEAEIQLNCLAQAGSDPTCTSAVDLWSLGIILLEMFGMKLKHIVRSQEWKANSALI 258
 DB 170 APEI-----LGSKYSTAVDWSLGCIFAEW-----VTRRALFPGDSE--I 209
 QY 259 DHIFASKAVVN-----AAIPAYHL-----RDLIKSMIHD 287
 DB 210 DQLFRIFRTLTGTPDEVVWPGVTSMDDYKPSFPKMARQDFSKVVPPLDEBDSLSLQMLHY 269
 QY 288 DSRRIIPAEALCSPPFSIPFAHIEDVLMPTPVRLNLVDDVYLGNBEEYEDVVEVY 347
 DB 270 DPKRIRISAKAALAHPPF-----QD-VTKPVPHRLMD-----PVATWV 307
 QY 348 KERCOKYGPVSLVLEKENGPGQVPEVYANAGDSKAAOKLITGRMFDEKFFVAT 402
 DB 308 SKGEELFTGVVPILVELDGDVNGHKF-SVSGESGDRYTGKLT-----LKFICTT 356

RESULT 13

ADD21398 standard; protein; 298 AA.

AC ADD21398;

DT 15-JAN-2004 (first entry)

DE Golden hamster cyclin dependent kinase (cdk) 2 protein.

KW continual growth; cultured cell; cyclin dependent kinase; cdk4; cdk2;
 KW cdk6; activating mutation; cell growth; cell division; cell cycle;
 KW cancer-causing agent; continual growth-induced cell; enzyme;
 KW golden hamster.
 XX Mesocricetus auratus.
 XX WO2003044169-A2.
 XX 30-MAY-2003.
 XX 15-NOV-2002; 2002WO-US036729.
 XX 15-NOV-2001; 2001US-0334760P.
 XX (UTEM) UNIV TEMPLE.
 XX Reddy PE, Rane SG, Mettus RV;
 XX WPI; 2003-449813/42.
 XX N-PSDB; ADD21399.
 XX A composition for reversibly inducing continual growth in normal cells
 PT comprises a cyclin dependent kinase protein (e.g. cdk4, cdk2 or cdk6) or
 PT its active fragment, derivative, homolog or analog, having an activating
 PT mutation.
 XX Disclosure; Page 104; 77pp; English.

CC This invention relates to a novel composition for inducing a reversible
 CC state of a continual growth in cultured cells and comprises at least one
 CC compound comprising a cyclin dependent kinase (cdk4, cdk2 or cdk6
 CC protein having an activating mutation. Growth and division of living
 CC cells involve a regular series of events and processes that comprise the
 CC cell cycle. Cyclin dependent kinases cdk2, cdk4 and cdk6 are involved in
 CC the control of G1, the point at which cells irreversibly commit to DNA
 CC synthesis and thus enter the cell cycle. The invention is useful in
 CC reversibly inducing continual growth in normal cells and may allow the
 CC screening of cancer-causing agents with the continual growth-induced
 CC cells. The present sequence is that of the golden hamster cdk2 protein
 CC related to the invention. Note: Due to an error in the specification or
 CC sequence listing, the Seq ID numbers given in the disclosure do not
 CC correspond to those given in the sequence listing. It is therefore
 CC unclear which Seq ID number corresponds to which sequence and exactly
 CC which sequence is being claimed.

XX Sequence 298 AA;

Query Match 10.7%; Score 235; DB 7; Length 298;
 Best Local Similarity 29.0%; Pred. No. 3.7e-16;
 Matches 98; Conservative 45; Mismatches 117; Indels 78; Gaps 16;

QY 23 MVOGSLGSGSSASVYRVRCGNGSPGALRQF-LPFGTGAASAAYGPKERAALE 81
 DB 4 FQKVEKIGEGTGVVYKAK--NKLGEVVALKKIRLDTEBVPSTAL-----REISLTK 56
 QY 82 QLOGHNRITVLYGVFTIHSPNVPBSRCLLELDVSVSELLYSSHOGCSMMMIQHCARD 141
 DB 57 EL-NHNPNIYKLDLV--IHTEKTL---YLVEFLHQDLKKFMDSALVGTIPLPIKSYLQ 110
 QY 142 VLEALFLHHEGVYADLKPRNIIWSAENECFKLIDFGLSFKGNQDVKY---IQTDGVR 198
 DB 111 LLOGLAFLCHSHRVLRDLKPNQNLINTEG-AIKLADFGIARAFGVPRVYTHEVVTLMYR 169
 QY 199 APEAEIQLNCLAQAGSDPTCTSAVDLWSLGIILLEMFGMKLKHIVRSQEWKANSALI 258
 DB 170 APEI-----LGSKYSTAVDWSLGCIFAEW-----VTRRALFPGDSE--I 209
 QY 259 DHIFASKAVVN-----AAIPAYHL-----RDLIKSMIHD 287
 DB 210 DQLFRIFRTLTGTPDEVVWPGVTSMDDYKPSFPKMARQDFSKVVPPLDEBDSLSLQMLHY 269
 QY 288 DSRRIIPAEALCSPPFSIPFAHIEDVLMPTPVRLNLVDDVYLGNBEEYEDVVEVY 347

Db 270 DEWKISAKALAHPEF-----OD-VTKPVPHRL 298

RESULT 14
AAB19094
ID AAB19094 standard; protein; 433 AA.
XX AAB19094;
XX
DT 19-FEB-2001 (first entry)
XX
DE Protein encoded by reporter gene YKL161C from protein kinase C pathway.
XX
KM Reporter gene; ergosterol pathway; protein kinase C pathway;
XX invasive growth pathway; biological pathway.
XX
OS Saccharomyces cerevisiae.
XX
PN WO20058520-A1.
XX
PD 05-OCT-2000.
XX
PF 29-MAR-2000; 2000WO-US008555.
XX
PR 31-MAR-1999; 99US-00282243.
XX
PA (ROSE-) ROSETTA INPHARMATICS INC.
XX
PI Roberts CJ;
XX
DR WPI; 2000-664929/64.
XX
DR N-PSDB; AAA86704.
XX
PT Identifying reporter and target genes for particular biological pathway
PT of interest, useful for drug designing, involves identifying a gene which
PT clusters to geneset associated with biological pathway.
XX
PS Disclosure; Fig 20; 239p; English.
XX
CC The present sequence is encoded by a reporter gene from the protein
CC kinase C pathway. The specification also describes reporter genes from
CC the ergosterol pathway, and the invasive growth pathway. The genes are
CC identified using methods of the invention. The specification describes
CC methods for identifying a reporter gene or a target gene for a particular
CC biological pathway in a cell. The methods comprise identifying a gene
CC which clusters to a geneset associated with the pathway. The gene
CC identified is necessary for normal function of the pathway. The methods
CC are useful for analysing the activity of biological pathways and in drug
CC designing, drug therapies, or other biological agents e.g. insecticides,
CC herbicides, fungicides, antibodies or antivirals. The reporter gene is
CC useful as a specific target for drugs which are designed to enhance,
CC inhibit or modulate a particular biological pathway
XX
SQ Sequence 433 AA:
Query Match 10.6%; Score 234.5; DB 3; Length 433;
Best Local Similarity 29.5%; Pred. No. 7.4e-16;
Matches 89; Conservative 45; Mismatches 97; Indels 71; Gaps 16;
QY 75 KERAALBQLOGHNIYVTVGVFTIHFSNVPSRCLL--ELLDVSVSELLYSSHQCSM 132
DB 71 RELKLRHRLHGHNIYVTVGVFTIHFSNVPSRCLL--ELLDVSVSELLYSSHQCSM 127
QY 133 WMIQHCARDVLEALFLAHGCVYHADLKPRNIIWSAENEC-FKLIDFGLS-----FKEG 185
DB 128 AHFQSFYQILCALKTYHSANVLAHCDKPKNLL--VNSDQQLKCNFGJSCSYSENHKVN 185
QY 186 NQGVK-YIQDGRAPAEALQNLQAGLOSDTECSAVDNLGLITLLE-----MFGS 238
DB 186 DGIKIGTISIMWKAPI-----LLNYOECTKAVDIWSTGCILLELLGRKPMFEG 235
QY 239 M-----KLKHTVR-----SQEWKNSSAIIDHIF-----ASKAVVNAIDPAYHL 277

Db 236 KDYVDHLNIIQLIGTPEETLQE--IASQKYNYIFQGNIPGRSFESILPGANP--EA 291

QY 278 RDLIKSMHDDPSRRIRPAMALCSFPFSI-----PFA-----PHIEDLVWLPTPV 322
DB 292 IELTKRMLEFPDKRITVEDALEHPYLSMWHMDIDEEFSCQKTFRPFETHISMALGNEV 351

QY 323 LR 324
DB 352 IK 353

RESULT 15
ADP87472
ID ADP87472 standard; protein; 433 AA.
XX ADP87472;
XX
DT 23-SEP-2004 (first entry)
XX
DE S cerevisiae glucan synthase pathway protein YKL161C SegID14.
XX
KM glucan synthase pathway; RNA expression; protein expression; YOL1 13W;
KM SKM1; YNR066C; YLR121C; YPS3; YHR209W; YKL161C; YFR030W; FLO1; YFR030W;
KM MET10; antifungal agent.
XX
OS Saccharomyces cerevisiae.
XX
PN WO2004057033-A1.
XX
PD 08-JUL-2004.
XX
PF 17-DEC-2003; 2003WO-US040532.
XX
PR 19-DEC-2002; 2002US-00324035.
XX
PA (ROSE-) ROSETTA INPHARMATICS LLC.
XX
PI Phillips JW;
XX
DR WPI; 2004-500308/47.
XX
DR N-PSDB; ADP87471.
XX
PT Determining whether a molecule affects the function or activity of a
PT glucan synthase pathway in a S. cerevisiae cell by determining a change
PT in the RNA expression or protein expression in the cell of at least one
PT target polynucleotide.
XX
PS Disclosure; SEQ ID NO 14; 132p; English.
XX
CC This invention relates to a novel method of determining if a molecule
CC affects the function or activity of a glucan synthase pathway in a
CC Saccharomyces cerevisiae cell which comprises determining if the RNA
CC expression or protein expression in the cell of at least one target
CC polynucleotide sequence is changed relative to the expression of said
CC target polynucleotide sequence in the absence of the molecule. The target
CC polynucleotide sequence is selected from the group consisting of YOL1 13W
CC (SKM1), YNR066C, YLR121C (YPS3), YHR209W, YKL161C, YFR030W (FLO1) and
CC YFR030W (MET10). The method is useful for determining whether a molecule
CC affects the function or activity of a glucan synthase pathway in an S
CC cerevisiae cell, possibly allowing development of antifungal agents for
CC use against a variety of pathogens. The present sequence is that of the
CC protein encoded by an S cerevisiae gene which may be used in the method
CC of the invention.
XX
SQ Sequence 433 AA:
Query Match 10.6%; Score 234.5; DB 8; Length 433;
Best Local Similarity 29.5%; Pred. No. 7.4e-16;
Matches 89; Conservative 45; Mismatches 97; Indels 71; Gaps 16;
QY 75 KERAALBQLOGHNIYVTVGVFTIHFSNVPSRCLL--ELLDVSVSELLYSSHQCSM 132

Db 71 RELKLLRHLRGHPNIVLFDPT-DIVFYPNGALNGVYLYBELMECDLSQII--RSEQRLD 127
QY 133 WMIOHCARDVLEALAFIHHGVVHADLKPPNIIIMSAENEC-FKLIDFGLS-----FKEG 185
Db 128 AHQSPFIYQILCALKYIHSANVHCDLKPKNLL--VNSDQOLKICNFGLSCSYSENHKYN 185
QY 186 NODVK-YIQTDGYRAPEAEIQNCLAQAGLQSDTECTSAVDLMSGLIILLE-----MSG 238
Db 186 DGFIKGYITSIWYKAPDEI-----LITNYQECTKAVDIWSTGCILAEILGRKPMFEG 235
QY 239 M-----KLKHTVR-----SOEMKANSALIDHTF-----ASKAVVNNAIPAYHL 277
Db 236 KDYVDHLNHLQILGTPEETLOE--IASQKVYNYLFQGNIPGRSFESTILPGANP--EA 291
QY 278 RDLIKSMLHDDPSRRIPAEMALCSPFESI-----PFA-----PHIEDLVMLPTPV 322
Db 292 LELUKMLERDPKKRITVEDALEHPYLSMWHIDIDEEFSCQKTRPFEFHIESMAELGNEV 351
QY 323 LR 324
Db 352 IK 353

Search completed: November 29, 2004, 11:30:38
Job time : 111.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OW protein - nucleic search, using frame_plus_p2n model

Run on: November 29, 2004, 11:25:02 ; Search time 4563 Seconds
(without alignments)
4342.403 Million cell updates/sec

Title: US-10-798-532-4

Perfect score: 2206
Sequence: 1 MAGSCAGMABPPRFLEARG.....VATFYPLSAVKRGVLYQTLL 419

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame_plus_p2n.model -DEV=xlp
-e=/cgn2.1/USPTO.spool_p/US10798532/runat_29112004_112452_15890/app_query.faeta_1.1166
-DB=GenEmbl -QFMT=fastap -SUFFIX=xge -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bites -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10798532.@CNCN_1_1_6317@runat_29112004_112452_15890 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONELIG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:.*
1: gb_ba:.*
2: gb_hcg:.*
3: gb_in:.*
4: gb_om:.*
5: gb_ov:.*
6: gb_pac:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_scs:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vl:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2195	99.5	1260	6	CQ715025 Sequence
2	2195	99.5	1260	6	AX262521 Sequence
3	2195	99.5	2598	6	AX262519 Sequence
4	2195	99.5	2909	9	HSAS36197 Homo sapi

5	2195	99.5	4065	6	CQ412057	CQ412057 Sequence
6	2195	99.5	4065	6	AX188296	AX188296 Sequence
7	2188	99.2	1260	6	AX166543	AX166543 Sequence
8	2188	99.2	1824	6	AX680136	AX680136 Sequence
9	2183	99.0	1633	10	RNKISRNA	RNKISRNA
10	2183	99.0	3244	10	RNU70372	RNU70372
11	2182	98.9	2041	10	BC058732	BC058732 Mus muscu
12	2174	98.5	110000	2	AC129141_1	Continuation (2 of
13	2157	97.8	1254	10	MMPKIS	Y10725 Mus musculu
14	1990	90.2	1367	9	BC026046	BC026046 Homo sapi
15	1890.5	85.7	2340	9	HSW05327	ALB34136 Homo sapi
16	1761.5	79.9	1157	5	CR386837	CR386837 Gallus ga
17	1721	78.0	1135	9	AK058195	AK058195 Homo sapi
18	1634	74.1	2041	9	BC014917	BC014917 Homo sapi
19	1050	47.6	259562	2	AC094499	AC094499 Rattus no
20	1045	47.4	601	10	MMKISRNA	X82320 M.musculus
21	905	41.0	105902	9	ALJ59699	ALJ59699 Human DNA
22	905	41.0	167995	2	AC040901	AC040901 Homo sapi
23	901	40.8	211812	2	AC027205	AC027205 Homo sapi
24	802.5	36.4	211635	2	AC112763	AC112763 Rattus no
25	789	35.8	529	6	AX186343	AX186343 Sequence
26	766.5	34.7	722	6	BD221413	BD221413 Human gen
27	729	33.0	470	6	AX187546	AX187546 Sequence
28	721	32.7	100207	10	AX180177	AX180177 Mus muscu
29	694.5	31.5	200095	10	AC121606	AC121606 Mus muscu
30	694.5	31.5	220806	10	AC123650	AC123650 Mus muscu
31	663	30.1	462	6	AX184680	AX184680 Sequence
32	366	16.6	60640	10	AX180176	AX180176 Mus muscu
33	366	16.6	185810	2	AC136052	AC136052 Rattus no
34	361	16.4	231	6	BD220137	BD220137 Human gen
35	327.5	14.8	205752	2	AC119431	AC119431 Mus muscu
36	311.5	14.1	419	6	CQ393016	CQ393016 Sequence
37	311.5	14.1	419	6	CQ399411	CQ399411 Sequence
38	267	12.1	185810	2	AC136052	AC136052 Rattus no
39	250.5	11.5	110000	8	CR382128_04	Continuation (5 of
40	250.5	11.4	1269	10	AB096078	AB096078 Mus muscu
41	250.5	11.4	1395	10	AB096079	AB096079 Mus muscu
42	249	11.3	1065	10	HAMCDK2A	D17350 Mesocricetu
43	248.5	11.3	2809	3	AY064464	AY064464 Caenorhab
44	246.5	11.2	1400	10	RATCDK2A	D28753 Rattus ratt
45	246.5	11.2	2250	10	BC061832	BC061832 Rattus no

ALIGNMENTS

RESULT 1
LOCUS CQ715025 1260 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 959 from Patent WO02068579.
ACCESSION CQ715025
VERSION CQ715025.1 GI:42275882
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kites, such as nucleic acid arrays, comprising a majority of
humanecons or transcripts, for detecting expression and other uses
thereof

JOURNAL Patent: WO 02068579-A 959 06-SEP-2002;
PB Corporation (NY) (US)

FEATURES
source Location/Qualifiers
1..1260
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores: 2.07e-188 Length: 1260
Pred. No.: 2195.00 Matches: 417
Score:

Percent Similarity: 99.76% Conservative: 1
 Best Local Similarity: 99.52% Mismatches: 1
 Query Match: 99.50% Indels: 0
 DB: 6 Gaps: 0

US-10-798-532-4 (1-419) x CQ15025 (1-1260)

```

QY 1 MetAlaGlySerGlyCySaLaTPGlyAlaGluProProArgPheLeuGluAlaPheGly 20
DB 1 ATGGCGGGAATCCGGCTGCGCTGGGGGCGGAGCCGCGCTTTCTGAGGGCTTCGGG 60
QY 21 ArgLeuTPGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
DB 61 CGGCTGTGGGAGTACAGAACCGCTGTGGTACGGGCTCCCGCTCGGGTATCGGGTT 120
QY 41 ArgCyGSerGlyAAsnProGlySerProProGlyValAlaLeuArgGlnPheLeuProProGly 60
DB 121 CGCTGTGCGGCAACCTGTGCTGCGCCCGCGGCTCCAGCAAGCTTCTTGGCCGCA 180
QY 61 ThrThGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgGlyArgAlaAlaLeu 80
DB 181 ACCACCGGGGCTGCGGCTGTGCGCGCAATAGTTTCCGCAAGAGAGGGGCGGCTG 240
QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
DB 241 GAACAGTTGCAGGGGTACAGAAACATCGTACCTTGTATGAGAGTGTTCATCAATCACTTT 300
QY 101 SerProAsnValProSerArgCySLeuLeuLeuGluLeuLeuAspValSerValSerGly 120
DB 301 TCTCCAAATGTGCATCACCGCTGTCTGTCTTGAACCTCGATGTCAGTGTTCGGAA 360
QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCySerMetTyrMetIleGlnHisCySaLaArg 140
DB 361 TTGCTCTTATATTCAGTCCACGAGGTGTTCATGTGATGATGATACAGCATTTGCCGA 420
QY 141 AspValLeuGluAlaLeuAlaPheLeuHisHisGlnGlyTyrValHisAlaAspLeuVal 160
DB 421 GAGTGTGTGGAGGCCCTTGTCTTTCATCATGAGAGGCTATGTCCATGCGGACCTCA 480
QY 161 ProArgAsnIleLeuTyrSerAlaGluAsnGlyCyAspPheLeuIleAspPheGlyLeu 180
DB 481 CCACGTACATATTTGAGAGTGCAGAGATGAATGTTTAACTCATGTGACTTGGACTT 540
QY 181 SerPheLeuGlnGlyAsnGlnAspValTyrTyrIleGlnThrAspGlyTyrArgAlaPro 200
DB 541 AGCTTCAAAGAGGCAATCAGATGTAAAGTATATTCAGACAGCGGGTATCGGCTCCA 600
QY 201 GluAlaGluLeuGlnIleAsnCySLeuAlaGlnAlaGlyLeuGlnSerAspThrGlyCThr 220
DB 601 GAAGCAGAAATTCGAAATTCCTGGCCGAGGCTGCGCTGCAGAGTATACAGAAATGTA 660
QY 221 SerAlaValaAspLeuTyrSerLeuGlyIleIleLeuLeuGlnLeuPheSerGlyMetLeu 240
DB 661 TCGAGCTTATCTGTGAGACCTTAGAATCATTTTACTGAAATTTCTTCAGAAATGA 720
QY 241 LeuLeuHisThrValArgSerGlnGluTyrPlyAlaAsnSerSerAlaIleIleAspHis 260
DB 721 CTGAACATATCAGATCAGATCTCAGAAATGAGAGCAACAGTTCTGTATATTAATGATCAC 780
QY 261 IlePheAlaSerIleValAlaValaAsnAlaAlaIleProAlaTyrHisPheLeuAspLeu 280
DB 781 ATATTTGGCCAGTAAAGAGGTGTAATGCGCAATTCACACCTATACCTTAAAGACCTT 840
QY 281 IleLeuSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCyS 300
DB 841 ATCAAAAGCATGCTCTCTATGATTCACAGCAAGAAATTCCTGCTGAATAGCATTTGTGC 900
QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
DB 901 AGCCCATTTCTTAGCATTTCTTTTGGCCCTCATATTAAGATCTGGCATGCTTCCACT 960
QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspArgTyrLeuGlnGlnGluGlyTyr 340
  
```

DB 961 CCAGTCTCAAGACTGTCTGAATGTCTGATGATATCTTGAATCAAGAGCAATAT 1020

QY 341 GluAspValaGluAspValaGlyGlnGlyCySglnTyrGlyProValaSerLeu 360

DB 1021 GAAGATGTGTGAAGATGTAAAGAGAGTGTCAAAAATATGACAGGTGTATCTCTTA 1080

QY 361 LeuValProGlyGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380

DB 1081 CTGTGTCCAAAGAAATCTGTGCGAGAGCAAGTCTTGTGTGATGTGCAATGCTGTG 1140

QY 381 AspSerValaAlaGlnIleLeuLeuThrGlyArgMetPheAspGlyAspPheVala 400

DB 1141 GATTCCAAAGCTGCCAGAAATTACTGACTGAAAGATTTGATGGAAATTTGTGTG 1200

QY 401 AlaThrPheTyrProLeuSerAlaTyrTyrAsArgGlyTyrLeuTyrGlnThrLeuLeu 419

DB 1201 GCTACATTTACCCGTGATGCTGCTTCAAGAGGGATATCTGTATCAACCTTGCTT 1257

RESULT 2

AX262521 1260 bp DNA linear PAT 26-OCT-2001

LOCUS Sequence 12 from Patent WO0173050.

DEFINITION AX262521

ACCESSION AX262521

VERSION AX262521.1 GI:16511408

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 Meyers, R. 3714, 16742, 23546, and 13887 novel protein kinase molecules and uses therefor Patent: WO 0173050-A 12 04-OCT-2001; Millennium Pharmaceuticals, Inc. (US)

JOURNAL

FEATURES

source 1..1260

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 2,07e-188 Length: 1260

Score: 2195.00 Matches: 417

Percent Similarity: 99.76% Conservative: 1

Best Local Similarity: 99.52% Mismatches: 0

Query Match: 99.50% Indels: 0

DB: 6 Gaps: 0

US-10-798-532-4 (1-419) x AX262521 (1-1260)

```

QY 1 MetAlaGlySerGlyCySaLaTPGlyAlaGluProProArgPheLeuGluAlaPheGly 20
DB 1 ATGGCGGGAATCCGGCTGCGCTGGGGGCGGAGCCGCGCTTTCTGAGGGCTTCGGG 60
QY 21 ArgLeuTPGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
DB 61 CGGCTGTGGGAGTACAGAACCGCTGTGGTACGGGCTCCCGCTCGGGTATCGGGTT 120
QY 41 ArgCyGSerGlyAAsnProGlySerProProGlyValAlaLeuArgGlnPheLeuProProGly 60
DB 121 CGCTGTGCGGCAACCTGTGCTGCGCCCGCGGCTCCAGCAAGCTTCTTGGCCGCA 180
QY 61 ThrThGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgGlyArgAlaAlaLeu 80
DB 181 ACCACCGGGGCTGCGGCTGTGCGCGCAATAGTTTCCGCAAGAGAGGGGCGGCTG 240
QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
DB 241 GAACAGTTGCAGGGGTACAGAAACATCGTACCTTGTATGAGAGTGTTCATCAATCACTTT 300
QY 101 SerProAsnValProSerArgCySLeuLeuLeuGluLeuLeuAspValSerValSerGly 120
  
```

```
Db 301 TCTCCAAATGGCCATCAGCTGTCTGTGGCTTGAACCTCTGATGTCAGTGTTCGGA 360
Qy 121 LeuLeuLeuTySerSerHisGlnGlyCySerMetTrpMetIleGlnHisGValAlaArg 140
Db 361 TTGCTCTTATATTCAGTCAACCGGGTGTTCATGTGATGATGATGATGATGATGATGATG 420
Qy 141 AspValLeuGlnAlaLeuAlaPheLeuHisHisGlnGlyTyrTyrAlaHisAlaAspLeuVal 160
Db 421 GATGTTTGGAGGCGCTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 480
Qy 161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCyPheLeuLeuIleAspPheGlyLeu 180
Db 481 CACCGTACATATTTGGAGTGCAGAGTGCAGAGATGTTTAACTCATTTGACTTGGACTT 540
Qy 181 SerPheLeuGlnGlyValangInAspValTyrTyrIleGlnThrAspGlyTyrAlaAlaPro 200
Db 541 AGCTTCAAAGAGAGGCATCAGATGTAAGTATATTCAGACAGACGGGTATCCGGGCTCCA 600
Qy 201 GluAlaGlnLeuGlnAsnCySerLeuAlaGlnAlaGlyLeuGlnSerAspTrpGlyCySerThr 220
Db 601 GAAAGCAATTTGCAAAATTTGCTTGCCCGACGGCTGCGCTGCAAGTATACAGATGTAAC 660
Qy 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnMetPheSerGlyMetCys 240
Db 661 TCAGCTGTGATCTGTGAGAGCTTAAAGATCATTTTACTGGAATGTTTCTCAGGAATGAAA 720
Qy 241 LeuValHisThrValArgSerGlnGluTrpValAlaAsnSerSerAlaIleIleAspHis 260
Db 721 CTGAAACATATTCAGTCAATCTCAGGAATGGAAGCAACAGTCTCTATATATATGATCAC 780
Qy 261 IlePheAlaSerTrpValAlaValAlaAlaIleProAlaTrpHisIleuAspAspLeu 280
Db 781 AATTTTCCAGTAAGAGATGTTGTAATGCGCAATTCACGCTATACCTTAAGAGACTT 840
Qy 281 IleLeuSerMetLeuHisAspAspProSerArgArgIleProAlaGlnMetAlaLeuCyS 300
Db 841 ATCAAAAGCATGCTTCTATGATGATCCAGACAGAAAGATTCCTCTGAATGGCAATTGTGC 900
Qy 301 SerProPhePheSerIleProPheAlaProHisIleGlnAspLeuValMetLeuProThr 320
Db 901 AGCCCATTTCTTATGATCTTCTTGTGCCCCCTCATATTAAGATCTGGTCAATGCTTCCACT 960
Qy 321 ProValLeuArgLeuLeuAsnValIleAspAspAspTrpLeuGlnValangInGluGlyTyr 340
Db 961 CCAATCTTAAGATGCTGTAATGTCGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Qy 341 GluAspValAlaGlnAspValIleValGlnGluGlyCysGlnIleTyrIleProValIleSerLeu 360
Db 1021 GAAAGATGTTGTGAAGATGTAAGAAAGAGATGTCAGAAATATGAGACCAAGGTATCTCTA 1080
Qy 361 LeuValProIleValGlnAspValIleValGlnGluGlyValIlePheValGlnTyrAlaAlaAlaGly 380
Db 1081 CTTGTTCCTCAAGAGAAATCTCTGCAAGGACCAAGTCTTGTGATGATGATGATGATGATGAT 1140
Qy 381 AspSerTrpAlaAlaGlnIleValLeuLeuThrGlyArgMetPheAspGlyValPheValIle 400
Db 1141 GATTTCAAAGCTCGGAGAAATTAATGATGTAAGAGATGTTGATGAGGAAAGTTGTGTG 1200
Qy 401 AlaThrPheTyrProLeuSerAlaTyrIleValArgGlyTyrLeuTyrGlnThrLeuLeu 419
Db 1201 GCTACATTTCTACCGCTGAGTGCCTACAGAGGGGATCTGATCAAACTTGCTT 1257

RESULT 3
AX262519 2598 bp DNA linear PAT 26-OCT-2001
LOCUS DEFINITION Sequence 10 from Patent WO0173050.
ACCESSION AX262519
VERSION AX262519.1 GI:16511406
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
REFERENCE
1 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS Meyers, R.
TITLE 3714, 16742, 23546, and 13887 novel protein kinase molecules and
uses therefor
JOURNAL Patent: WO 0173050-A 10 04-OCT-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source location/Qualifiers
1..2598
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
268..1527
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD10238.1"
/db_xref="GI:16511407"
/translation="NAGSGCAMGAPRFLFAPGLMOWQSRSLGSSASVYRACCG
NPGSPGALKQFLPGTTGAASAASAEYFPRKRALLEOLQGRINIVTLVGVETIHPS
NVPSCRLLBELLDVSVSELLYSHQGSMWIOHCARDVLEALFLHHEGVHADLK
PNTILMSAENRCFKLIDPGLSPKGNODVXYIOTDGVAPRAEAIQNCILAQGLSDTE
CTSAYDLSLGIILLMFSGMKLKRTVRSQERKANSALIDHIFASKAVVNAIATYH
LRDLIKSMLEHDPSSRIIPAEVALCSPPFSIIPAPHIDLVMLPTVLELVNLDDYL
ENBEYEDVDEYKEEQYGPVSLVLPKENPGRGVFEVYANAGDSKAQKLTGR
MDFKRVATFYPLSAVRYLYQTLL"
ORIGIN
Alignment Scores:
Pred. No.: 5.7e-188 Length: 2598
Score: 2195.00 Matches: 417
Percent Similarity: 99.76% Conservative: 1
Best Local Similarity: 99.52% Mismatches: 1
Query Match: 99.50% Indels: 0
Gaps: 0
US-10-798-532-4 (1-419) x AX262519 (1-2598)
Qy 1 MetAlaGlnSerGlnCySerAlaTrpGlyAlaGluProProArgPheLeuGlnAlaPheGly 20
Db 268 ATGGCCGAGATCCGGCTGCGCCCTGCGGCGCGAGCCGCCGCTTCTTGGAGGCTTCGGG 327
Qy 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
Db 328 CGGCTGTGCAAGTACAGAGCCGTCTGGGTACCGGCTCTCCGCTCGGTATGCGGTT 387
Qy 41 ArgCySerGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
Db 388 CCTGCTGCGGCAACCTGCTGCGCCCGCGCGCTTCACACAGTCTTGTGCGGACAGA 447
Qy 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgIleGlnArgAlaAlaLeu 80
Db 448 ACCACCGGGGCTGCGGCTCTCTCCCGCGAGTATGGTTCCGAAAGAGGGCGGCGCTG 507
Qy 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
Db 508 GAAAGTTGCAAGGTACAGAAACATCGTGACTTTGTATGAGTGTATTAATCAATCACTTT 567
Qy 101 SerProAsnValProSerArgCySerLeuLeuLeuGlnLeuAspValSerValSerGlu 120
Db 568 TCTCCAAAGTGCATACAGCTGTCTGTGCTTGAACCTCTGATGATGATGATGATGATGAT 627
Qy 121 LeuLeuLeuTySerSerHisGlnGlyCySerMetTrpMetIleGlnHisGValAlaArg 140
Db 628 TTGCTCTTATATTCAGTCAACCGGGTGTTCATGTGATGATGATGATGATGATGATGATGAT 687
Qy 141 AspValLeuGlnAlaLeuAlaPheLeuHisHisGlnGlyTyrTyrAlaHisAlaAspLeuVal 160
Db 688 GATGTTTGGAGGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 747
Qy 161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCyPheLeuLeuIleAspPheGlyLeu 180
Db 748 CACCGTACATATTTGGAGTGCAGAGTGCAGAGATGTTTAACTCATTTGACTTGGACTT 807
```

```

QY 181 SerPheLeuGluGlyAsnGlnAaspValIyPtyrIleGlnThrApgIYTYrAgaIapro 200
Db 808 AGCTTCAGAGAGGACATCAGATGTAAAGTATTTCAGACAGCGGATTCGGCTCCA 867
QY 201 GUAAGGluLeuGlnAenCySleuAlaGlnAlaGlyLeuGlnSerApthrGluCyPthr 220
Db 868 GAAGCAGAAATGGAAAATGGCTTGGCCGAGCTGGCTGAGATACAGAAATGATACC 927
QY 221 SerAlValAaspLeuTrpSerLeuGlyIleIleLeuLeuGlnuEcPheSerGlyMeClys 240
Db 928 TCAGCTGTTCATCTGTGAGACCTAGAAATCATTTTACTGGAATGTTCTCAGAAATGAAA 987
QY 241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerAlaIleIleAphHis 260
Db 988 CTGAAACATACAGATCAGATCTCAGAAATGAGAAACAGCTTCGTCTATTATTGATAC 1047
QY 261 IlePheAlaSerIyValAlaValAlaAlaIleProAlaTyrtHisleuAgaApleu 280
Db 1048 ATATTTGGCCGTAAGCAGTGTGAATGCGCAATTCAGCTATACCTAGAGACCTT 1107
QY 281 IleLysSerMetLeuHisAaspProSerArgArgIleProAlaGluMeAlaLeuCys 300
Db 1108 ATCAAAAGCATGCTTCATGATGATCCAGACAGAAATTCCTCTGAAATGGCATTTGTC 1167
QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAaspLeuValMeLeuProThr 320
Db 1168 AGCCCATTTCTTATAGCATCTTTTGGCCCTCATATTGAAGATCTGGATGCTTCCCACT 1227
QY 321 ProValLeuArgLeuLeuAanValIleuAaspAaspTyrtleuGlyengluGluGly 340
Db 1228 CCGTCTTAAAGCTGCTGAATGTGCTGATGATGATTAATCTTAAGAAATGAAGGAATAT 1287
QY 341 GluAaspValValGluAaspValIySgluGluCySgluLysIyTrpValIleSerLeu 360
Db 1288 GAAGATCTTCTAAGATGTAAAGAGAGACTCTCAAAATATGACACAGATGATCTCTA 1347
QY 361 LeuValProLysGluAanProGlyArgGlyGlnValPheValGluTyrtAlaAanAlaGly 380
Db 1348 CTGTGTCCAAGGAAATCTCTGACAGAGCAAGTCTTGTGATGATGCAATGCTGGGT 1407
QY 381 AapSerIyAlaAlaGlnLysleuLeuThrGlyArgMetPheAaspGlyLysPheValIle 400
Db 1408 GATTCCAAAGCTCGCAGAAATTAATGACTGGAAGATGTTGATGAGAAATTTGTTG 1467
QY 401 AlaThrPheTyrtProLeuSerAlaTyrtLysArgGlyTyrtleuTyrtGlnThrLeuLeu 419
Db 1468 GCTACATTCACCGCTGAGTGCCTACAGAGGGGATATCTGATCAAACTTGCCT 1524

RESULT 4
HSAS36197 2909 bp mRNA linear PRI 11-JAN-2003
LOCUS HSAS36197
DEFINITION Homo sapiens mRNA for Kis protein.
ACCESSION AJ536197
VERSION AJ536197.1 GI:27657360
KEYWORDS Kis protein. (human)
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1. Bieche, I., Manceau, V., Curmi, P. A., Laurendon, I., Lechar, S.,
Ieroy, K., Vidaud, D., Sobel, A., and Maucuer, A.
Quantitative RT-PCR reveals a ubiquitous but preferentially neural
expression of the Kis gene in rat and human
unpublished
2. (bases 1 to 2909)
JOURNAL
REFERENCE
AUTHORS Maucuer, A.
TITLE Direct Submision
JOURNAL Submitted (08-JAN-2003) Maucuer A., U440, Inserm, IFM, 17 rue du
fer a moulin, 75005 Paris, FRANCE
FEATURES
source
1..2909
/organism="Homo sapiens"

```

```

/mol_type="mRNA"
/db_xref="RZPD:IMAGp98M213680"
/db_xref="RZPD:IMAGp98M213680"
/clone="1.M.A.G.E. Consortium CloneID 145548"
/ligase_type="2 pooled tumors (clear cell type)"
1..2909
/gene="KIS"
/gene="KIS"
CDS
146..1405
/gene="KIS"
/codon_start=1
/product="KIS protein"
/protein_id="CD60192.1"
/db_xref="GI:27657361"
/translation="MAGSGCAGWAPPPFLAFLGLWQVGRLSGSSASVYRRCG
NGSPGALKQFLPPTGTAASAEVEFRERARLEOLQGRNIVTLVGVFTIHFS
NPSRCLLELDIVSELLIYSSHQGSMNIOHCAVDLEALFLHBEYVHADLK
PNITLMSAENECFKLIDGLSPKEONOVKXIOGTGYRAREALONCLAOGLSDTE
CTSAVDNLSLGIILEMSKMKLKTATVASCOKANSASLIDHIFAPKLVNNAADPAH
LADLTKSLHDDPSRRITPAEMALCSPPSIFPAPHIDLVMLPFLVRLVAVDDYL
ENESEYEDVEDVEKCECKYGPVSLVLPKENPGGVFEVANAAGDSKAAQKLTGR
MDGKRVVATFPYPLSAVRGYLYQLL"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 6,68e-188 Length: 2909
Score: 2195.00 Matches: 417
Percent Similarity: 99.76% Conservative: 1
Best Local Similarity: 99.52% Mismatches: 1
Query Match: 99.50% Indels: 0
DB: 9 Gaps: 0
US-10-798-532-4 (1-419) x HSAS36197 (1-2909)

```

```

QY 1 MetAlaGlySerGlyCySAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
Db 146 ATGGCGGGAATCCGGCTGCGCCCTGGGGCCGAGACCCCGCTTCTGAGGCCCTTCCGG 205
QY 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrtArgVal 40
Db 206 CCGTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCTCGATGCGGGTT 265
QY 41 ArgCySLeuAanProGlySerProProGlyValAlaAanGlnPheLeuProProGly 60
Db 266 CGCTGTGGGCAACCTGTGCTGCGCCCGCGCTCAACAGATTTCTGCGCCAGGA 325
QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrtGlyPheArgGlyGluArgAlaAlaLeu 80
Db 326 ACCACCGGGGCTGCGGCTCTGCCCGAGTATGTTCCGAAAGAGGGCGGCGCTG 385
QY 81 GluGlnLeuGlnGlyHisArgAanIleValThrLeuTyrtGlyValPheThrIleHisPhe 100
Db 386 GAACAGTTGACAGGTCACAGAAACATCGACTTGTATGAGTGTTAACAATCCACTTT 445
QY 101 SerProAanValProSerArgCySleuLeuLeuGluLeuAaspValSerValSerGlu 120
Db 446 TCTCCAAATGTCCATCAGCTGTCTGTGAATCTCGAGATGACAGTGTTCGGA 505
QY 121 LeuLeuLeuTyrtSerSerHisGlnGlyCySerSerMetTrpMetIleGlnHisCySAlaArg 140
Db 506 TTGCTCTTATATTCAGATCACAAGGTTGTTCCATGTGATGATACAGCATTTGCCGA 565
QY 141 AapValLeuGluAlaLeuAlaPheLeuHisHisGlnGlyTyrtValHisAlaAaspLeu 160
Db 566 GATGTTTGAAGGCCCTTCTTCTTCATCATGAGGCTATGTCATCGGACCTCAAA 625
QY 161 ProArgAanIleLeuTrpSerAlaGluAengluCySphelysleuIleAapheGlyLeu 180
Db 626 CCACGTAAATATTTGTGAGTGCAGAGATGATATTTTAACTATGACATTTGACTT 685
QY 181 SerPheLeuGluGlyAlaGlnAaspValIyPtyrIleGlnThrAaspGlyTYrAgaIapro 200
Db 686 AGCTTCAGAGAGGACATCAGATGTAAAGTATATTTCAGACAGACGCGTATCGGCTCCA 745

```

QY 201 GIUAlaGluLeuGlnAnbCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
 DB 746 GAAGCAGAAATGCAAAATGCTTGCCCAAGCGCTGGCTGAGATGATACAGATATGAC 805
 QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnMetPheSerGlyMetLeu 240
 DB 806 TCAGCTGCTGATCTGAGAGCCCTAGGAATCATTTTCTGAGAAATGTTCTCAGAAATGAAA 865
 QY 241 LeuLeuHisThrValArgSerGlnGluTrpLeuAlaAsnSerSerAlaIleIleAspHis 260
 DB 866 CTGAACACATACATGATCTCAGAAATGAGAGCAACAGTTCCTGCTATTATTGATCAC 925
 QY 261 IlePheAlaSerIleValAlaValAlaAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
 DB 926 ATTTTGGCCAGTAAAGAGATGAGTGAATCCGCAATTCACCTTACCTTAAGACCTT 985
 QY 281 IleLeuSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
 DB 986 ATCAAAAGCATGCTTCATGATGATCCAGACAGAAATTCCTGCTGAAATGGCATTTGTC 1045
 QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
 DB 1046 AGCCCATCTCTTATGACATCTCTTTGCCCCCTCATATGAAATGATGCTGATGCTCCACT 1105
 QY 321 ProValLeuArgLeuLeuAnbValLeuAspAspAspTyrLeuGlnArgGluGluTyr 340
 DB 1106 CCAAGTCTAAGATCTGCTGATGCTGCTGATGATGATTTATCTTGAAGATGAAGAGATAT 1165
 QY 341 GluAspValAlaGluAspValAlaGluGluGluCysGlnLeuTyrGlyProValAlaSerLeu 360
 DB 1166 GAAGAGTGTAGAGATGTAAGAAAGAGAGATGTAAGAAATATGAGACAGATGATCTCTA 1225
 QY 361 LeuValProLeuGluAnbProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
 DB 1226 CTGTTCCTCAAGAGAAATCTCTGACAGAGCAAGTCTTGTGATGATGCAAAATGCTGCT 1285
 QY 381 AspSerIleAlaAlaGlnLeuLeuThrGlyArgMetPheAspGlyLeuPheValAla 400
 DB 1286 GATTCCAAGCTCGCAGAAATTAATCTACATGAAAGATGTTTATGGAGATTTGTTG 1345
 QY 401 AlaThrPheTyrProLeuSerAlaTyrLeuArgGlyTyrLeuTyrGlnThrLeuLeu 419
 DB 1346 GCTACATTTACCGCTGAGTGCCTACAGAGGGAGATATCTGATCAAACTTGCTT 1402
 RESULT 5
 CQ412057 4065 bp DNA linear PAT 23-JAN-2004
 LOCUS Sequence 19128 from Patent WO0170979.
 DEFINITION CQ412057
 ACCESSION CQ412057
 VERSION CQ412057.1 GI:41319838
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS Lee,J. and Lillie,J.
 TITLE Genes, compositions, kits, and method for identification,
 assessment, prevention, and therapy of ovarian cancer
 JOURNAL Patent: WO 0170979-A 19128 27-SEP-2001;
 Millennium Pharmaceuticals, Inc. (US)
 FEATURES
 source location/Qualifiers
 1..4065
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores: 1.07e-187 Length: 4065
 Pred. No.: 2195.00 Matches: 417
 Score: 99.76% Conservative: 1
 Best Local Similarity: 99.52% Mismatches: 1

Query Match: 99.50% Indels: 0
 DB: 6 Gaps: 0
 US-10-798-532-4 (1-419) x CQ412057 (1-4065)
 QY 1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProArgPheLeuGluAlaPheGly 20
 DB 268 ATGGCGGATCCGGCTGCGCTGGGGCGGAGCGCGCGCTTTCTGGAGCGCTTCGGG 327
 QY 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
 DB 328 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTACCGGCTCTCCGCTCGGTATGAGGTT 387
 QY 41 ArgCysCysGlyAspProGlySerProProGlyAlaLeuArgGlnPheLeuProGly 60
 DB 388 CGCTGCTGGCAGACCTTGCTGCTGCCCCCGCGCGCTCAACAGATCTTGGCGCAGGA 447
 QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLeuGluArgAlaAlaLeu 80
 DB 448 ACCACCGGGGCTGGCGCTCTGCGCGAGTATGTTCCGAAAGAGAGGGCGGCTG 507
 QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
 DB 508 GAACAGTTCACAGGTCAAGAAACATCGACTTGTATGAGTGTTCACATCCACTT 567
 QY 101 SerProAsnValProSerArgCysLeuLeuGluLeuLeuAspValSerValSerGlu 120
 DB 568 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAATCTTGATGATGATGTTTCGGA 627
 QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
 DB 628 TTGCTCTTATATTCACGATCACAGGGTGTTCATGTGATATACAGATTTGTCGCGA 687
 QY 141 AspValLeuGlnAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
 DB 688 GATGTTTGGAGGCCCTTGCTTCTTCATCATGAGGGCTATGTCATCGGACCTCAAA 747
 QY 161 ProArgAsnIleLeuTrpSerAlaGluAnbGluCysPheLeuLeuIleAspPheGlyLeu 180
 DB 748 CACGTAACATATTTGTGAGTGCAGAGAAATGATGTTTAACTCATTAATCTTGACTT 807
 QY 181 SerPheLeuGluGlyAsnGlnAspValIleTyrIleGlnThrAspGlyTyrArgAlaPro 200
 DB 808 AGCTTCAAAGAGGCAATCAGGATGTAATGATATTCAGACAGACGGGTATCGGGCTCA 867
 QY 201 GIUAlaGluLeuGlnAnbCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
 DB 868 GAAGCAGAAATGCAAAATGCTTGCGCCGAGCTGGCTGCAGAGATGATCAGAAATGATAC 927
 QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnMetPheSerGlyMetLeu 240
 DB 928 TCAGCTGTGATCTGAGAGCCCTAGGAATCATTTTAAGTAATGTTCTCAGGAATGAAA 987
 QY 241 LeuLeuHisThrValArgSerGlnGluTrpLeuAlaAsnSerSerAlaIleIleAspHis 260
 DB 988 CTGAACACATACATGATCTCAGAAATGAGAGCAACAGTTCGCTATTATTATGATCAC 1047
 QY 261 IlePheAlaSerIleValAlaValAlaAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
 DB 1048 ATTTTGGCCAGTAAAGAGATGAGTGAATCCGCAATTCACGCTTATCACTTAAGACCTT 1107
 QY 281 IleLeuSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
 DB 1108 ATCAAAAGCATGCTTCATATGATATCCAGACAGAAATTCCTGCTGAATGGCATTTGTC 1167
 QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
 DB 1168 AGCCCATCTTATGATGATCTCTTTGCCCCCTCATATTTGAAGATCTGCTATGCTCCACT 1227
 QY 321 ProValLeuArgLeuLeuAnbValLeuAspAspAspTyrLeuGlnArgGluGluTyr 340
 DB 1228 CCAAGTCTAAGATCTGCTGATGCTGATGATGATATCTTGAAGATGAAGAGATATAT 1287

QY	341	GlunbepvYalValGluabepValLyglunGlucyGlnlysrYrglyProvalValSerleu	360
Db	1288	GAAGATGTTGTAGAAAGATGTAAAGAAGAGTCAAAATATGACCAAGTGGATCTCTTA	1347
QY	361	LeuValProLybGluabnProGlyVArgslYglInValPheValGluTyPAlaAbnAlaGly	380
Db	1348	CTTGTTCAAAGGAAAAATCTCGGACGAGGACAAAGTCTTGTGTAGATGCAAACTGCGT	1407
QY	381	AhpSerlybAlaAlaGlnLylaleuLeuThrglyArgMetPheAhpGlyLybPheValVal	400
Db	1408	GATTCCAAAGCTGGCGCAGAAAATTAATCTGACTGGAGAGATGTTGAAGGAAAGTTTGTGTG	1467
QY	401	AlaThrPheTyrrProLeuSerAlaTyrrLybArgGlyTyrrLeuTyrglnThrLeuLeu	419
Db	1468	GCTACATTCTACCGCTGAGTGCTTACAAAGAGGGAGATATCTGTATCAAACTTGCTT	1524
RESULT 6			
LOCUS	AX188296	4065 bp	linear
DEFINITION	Sequence 3991 from Patent WO0142467.		PAT 06-AUG-2001
ACCESSION	AX188296		
VERSION	AX188296.1		
KEYWORDS	GI:15139769		
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
JOURNAL	Schlegel, R., Deeds, J., Berger, A. and Zhao, X.		
FEATURES	Genes, compositions, kits, and methods for identification, assessment, prevention, and therapy of cervical cancer		
source	Patent: WO 0142467-A 3991 14-VUN-2001; Millennium Predictive Medicine, Inc. (US)		
ORIGIN	Location/Qualifiers		
	1..4065		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
Alignment Scores:			
Pred. No.:	1,07e-187	Length:	4065
Score:	2195.00	Matches:	417
Percent Similarity:	99.76%	Conservative:	1
Best Local Similarity:	99.52%	Mismatches:	1
Query Match:	99.50%	Indels:	0
DB:	6	Gaps:	0
US-10-798-532-4 (1-419) x AX188296 (1-4065)			
QY	1	MetAlaGlySerGlyCyValaATrPglYAlaGluProPArgPheLeuGlnAlaPheGly	20
Db	268	ATGGCGGGATCCGGCTGGCGCTGGGGCGCGAGCGCGCGTTTCTGGAGGCTTGGGG	327
QY	21	ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyArgVal	40
Db	328	CGGCTGTGGCAGTACAGAGCCGTCTGGGTAGCGGCTCTCCGCTCCGTGTATGGGTT	387
QY	41	ArgCybCybGlyaaenProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly	60
Db	388	CGGTGCTCGGCAACCTCTGGCTCGCCCCCGCGCCCTCAAGCAATTTCTGGCCGACAGA	447
QY	61	ThrThrglyAlaAlaAlaSerAlaAlaGluTyrglyPheArgLybGluArgAlaAlaLeu	80
Db	448	ACCACCGGGGCTCGGCTCTGGCCGCGCAGTATGTTTCCGAAAGAGAAGGGCGGCTG	507
QY	81	GluGlnLeuGlnGlyHisArgAsnLleValThrLeuTyrglyValPheThrLleHisPhe	100
Db	508	GAAAGATTGCAGGGGCACAGAAACATCTGACTTGTGTATGAGAGTATTACAAATCACTTT	567
QY	101	SerProAnValProSerArgCybLeuLeuLeuGlnLeuLeuArgValSerValSerGln	120
Db	568	TCTCCAAATGTCACATCAAGCTGTCTGTGTTGTAATCTCGAGATGACAGTTCGGAA	627

Qy	121	LeuLeuLeuLeuTyrsSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisGlyValArg	140
Db	628	TYGGCTCTTAATATTCACATCCACGAGGTTGTTCCATGCGATGATACAGCATGTGGCCGA	687
Qy	141	AspValLeuGluAlaLeuAlaPheLeuHisHisGlnGlyTyrValHisAlaAspLeuLys	160
Db	688	GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATAGAGGCTATGTCCATGGGACCTCAA	747
Qy	161	ProArgAsnIleLeuTyrPserAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu	180
Db	748	CCAGGTAACTATATGTGGAGTGCAGAGAAAGAAATGTTTAAACTCATGTGACTTGGACTT	807
Qy	181	SerPheLysGlnGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro	200
Db	808	AGCTTCAAGAGAGGCATACAGAGATGAAGATATTCAGACACAGCGGGTATCGGGCTTCA	867
Qy	201	GluAlaGluLeuGlnAsnCysLysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr	220
Db	868	GAACCAAAATTCGAAATTCGCTTGCCACAGGCTGCGCTGCAGAGTGAATACAGATGATAC	927
Qy	221	SerAlaValAspLeuTyrPserLeuGlyIleIleLeuLeuGlnMetPheSerGlyMetLys	240
Db	928	TCAGCTCTTATCTGTGGAGCCTTAGAATCATATTTTACGAAATGTTCCTCAGGATATAAA	987
Qy	241	LeuLysHisThrValArgSerGlnGluTyrLysAlaAsnSerSerAlaIleIleAspHis	260
Db	988	CTGAACAACTACAGTCAGATCTCAGAAATGGAGGCACAAAGCTTCTGCTATATTGATCAC	1047
Qy	261	IlePheAlaSerLysValAlaValAlaAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu	280
Db	1048	ATATTTCCTCAATGAAGCAGTGGTGAATGGCCCAATTCACAGCTATCACTTAGAGACTT	1107
Qy	281	IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys	300
Db	1108	ATCAAAAGCATGCTTTCATGATGATCCAAAGAGAAAGATTCCTGCTGAATGGCATTTGTC	1167
Qy	301	SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr	320
Db	1168	AGCCCATTTCTTTAGCATTCCTTTTGGCCCCCTCATATTAAAGATCTGTCACTCCACT	1227
Qy	321	ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGluAsnGluGluTyr	340
Db	1228	CCAGTGTCTTAAGACTGCTGAGATGTGCTGATGATGATATATCTTGGAATGAAGAGAAATAT	1287
Qy	341	GluAspValValGluAspValLysGlnGluCysGlnLysTyrGlyProValValSerLeu	360
Db	1288	GAAGATTTTGTGAAGATGTGAAGAGAGAGCTCAAAAATATGGACCAAGTGGATATCTCTA	1347
Qy	361	LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly	380
Db	1348	CTTGCTCCAAAGGAAATTCCTGCGACAGAGACAAGTCTTTGTAGATATGCAAAATGCTGGT	1407
Qy	381	AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgGlyPheAspGlyLysPheValVal	400
Db	1408	GATTCCTCAAACTGCGCGAAGAAATTACTACTCGAAGAGATGTTTGTATGGGAATTTGTTGTG	1467
Qy	401	AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu	419
Db	1468	GCTACATTTCACCCGCTGAGTGCTTACAAAGGGGGAATATCTGATCAAACTTGCTT	1524
RESULT 7			
AX166543			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE			
1			
AX166543			
Sequence 34 from Patent MO0138503.			
AX166543			
AX166543.1 GI:14546888			
1260 bp DNA linear PAT 22-JUN-2001			

AUTHORS Plowman,G.D., Whyte,D., Manning,G.S., Sudarshanam,S.S., Martinez,R.,
 Flanagan,P. and Clary,D.S.
 TITLE Novel human protein kinases and protein kinase-like enzymes
 JOURNAL Patent: WO 0138503-A 34 31-MAY-2001;
 Sugen, Inc. (US)

FEATURES
 source Location/Qualifiers
 1.1260
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.:	8.88e-188	Length:	1260
Score:	2188.00	Matches:	416
Percent Similarity:	99.52%	Conservative:	1
Best Local Similarity:	99.28%	Mismatches:	2
Query Match:	99.18%	Indels:	0
DB:	6	Gaps:	0

US-10-798-532-4 (1-419) x AX166543 (1-1260)

QY 1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGlnAlaPheGly 20
 Db 1 ATGGCGGAGATCCGGCTGGCGCTGGGGGGCGAGCCGGCGGTTTCTGGAGGCGCTTCGGG 60
 QY 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTYrArgVal 40
 Db 61 CGGCTGTGGAGGTACAGAGCCGCTGGGTAGGGGCGCTCCCGCTCGGGTTCGGGTT 120
 QY 41 ArgCysGlyYanPProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
 Db 121 CGGTGCTGGCGCAACCTGCTGCCCCCGGGCGCTCAAGAGATTCTTGGCCGACGAA 180
 QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTYrGlyPheArgGlyGluArgAlaLeu 80
 Db 181 ACCACGGGAGCTCGGCGCTCTGCGCGCACTATGTTCCGCAAGAGAGGGCGGCTG 240
 QY 81 GluGlnLeuGlnGlyHisArgGlnIleValThrLeuTYrGlyValPheThrIleHisPhe 100
 Db 241 GAACAGTTGGAGGGTCAACAAACATCTGACTTGTATGAGATGTTTCAATCCACTT 300
 QY 101 SerProAsnValProSerArgCysLeuLeuLeuGlnLeuLeuAspValSerValSerGly 120
 Db 301 TCTCCAAATGTCACATCAACCTCTCTGTTCTTGAACCTCTGAGATGTCAGTTTCGGA 360
 QY 121 LeuLeuLeuTYrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
 Db 361 TTGCTCTTATATTCAGTCAACAGGGGTGTTTCATGTGATGATACAGCATTTGCCCCA 420
 QY 141 AspValLeuGlnAlaLeuAlaPheLeuHisGlnGlyTYrValHisAlaAspLeuLYs 160
 Db 421 GATGTTTGGAGGCCCTTGTCTTCTTCATCAAGAGGGGCTATCTCATGCGGACCTCAA 480
 QY 161 ProArgAsnIleLeuTrpSerAlaGluAsnGlnCysPheLeuLeuIleAspPheGlyLeu 180
 Db 481 CCAAGTAACATATTGGAGTGCAGAGATGAAATGTTTAAACTCATTTGAACCTT 540
 QY 181 SerPheLeuGlnGlyAsnGlnAspValLYrTYrIleGlnThrAspGlyTYrArgAlaPro 200
 Db 541 AGCTTCAAAGAGGCAATCAGATGTAAGTATATTCAAGCAACGGGTATCGGGCTCCA 600
 QY 201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
 Db 601 GAAGCAGAATTCGAAATTCCTTGGCCCAAGGCTGGCTGCAGATGATACAGATGATACC 660
 QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnMetPheSerGlyMetLYs 240
 Db 661 TCGAGCTTGATCTGGAGGCTTGAAGATCATTTTACGAAATGTTCTTCAGAAATGAAA 720
 QY 241 LeuLYsHisThrValArgSerGlnGluTrpLYsAlaAsnSerSerAlaIleIleAspHis 260
 Db 721 CTGAACATATCAGATCTCAGAAATGGAAGCAACAGTTCTGCTATATTATGATCAC 780

QY 261 IlePheAlaSerLYsAlaValValAsnAlaAlaIleProAlaTYrHisLeuArgAspLeu 280
 Db 781 ATATTTCGCAAGTAAGACATGCTGAATGCCGAAATTCAGCTTATCACTTAAGACACTT 840
 QY 281 IleLYsSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
 Db 841 ATCAAAAGATCTCTCATGATGATCCAGCAGAAATTCCTGCTGAATGCAATTGCTGCC 900
 QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
 Db 901 AGCCCATCTTTAGATCTCTTGGCCCTCATATTTGAAGATCTGTCATGCTCCCACT 960
 QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspArgTYrLeuGlyAsnGlnGluTYr 340
 Db 961 CCAAGTCTAAGACTGCTGAATGTGCTGATGATGATATCTTGAGAAATGAGAGAAATAT 1020
 QY 341 GluAspValAlaGluAspValLYsGlnGlyCysGlnLYsTYrGlyProValValSerLeu 360
 Db 1021 GAAGATGTTGTGAAGATGTAAAGAGAGTGTCAAAAATATGACACAGTGTATCTCTA 1080
 QY 361 LeuValProLYsGluAsnProGlyValArgGlyGlnValPheValGluTYrAlaAsnAlaGly 380
 Db 1081 CTGTGTCAAAGGGAATCTGCGACAGAGCAAGCTTTGTTGAGTATGCAATGCTGGT 1140
 QY 381 AspSerLYsAlaAlaGlnLYsLeuLeuThrArgYrArgMetPheAspGlyLYsPheVal 400
 Db 1141 GATTCCAAAGCTGCGCAGAAATTAAGTCTGAGAGAGATGTTGATGGAGAGTTTGTGG 1200
 QY 401 AlaThrPheTYrProLeuSerAlaTYrLYsArgGlyTYrLeuTYrGlnThrLeuLeu 419
 Db 1201 GCTACATTCTACCCCTGAGTGTCTTACAGAGAGGATATCTGTATCAACCTTGTCTT 1257

RESULT 8

AX680136 1824 bp DNA linear PAT 29-MAR-2003
 LOCUS Sequence 21 from Patent WO0181555.
 DEFINITION AX680136
 ACCESSION AX680136
 VERSION AX680136.1 GI:29369916
 KEYWORDS

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

AUTHORS

1 Gururajan,R., Patterson,C., Recipon,S.A., Tribouley,C.M.,
 Hafalia,A., Khan,F.Y., Yue,H., Au-Young,J., Bandman,O.,
 Baughn,M.R., Borowsky,M.L., Burford,N., Burrill,J.D., Elliot,V.S.,
 Gandhi,A.R., Kearney,L., Lal,P., Lu,D.A., Lu,Y., Tang,Y.T.,
 Zingler,K.A., Griffin,J.A., Hillman,J.J., Marcus,G.A., Nguyen,D.B.,
 Policky,J., Ramkumar,J., Thornon,M., Walla,N.K. and Walsh,R.T.
 Human Kinases

JOURNAL

Patent: WO 0181555-A 21 01-NOV-2001;
 Incyte Genomics, Inc. (US)

FEATURES

source Location/Qualifiers
 1.1824
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 3107952CB1"

ORIGIN

Alignment Scores:

Pred. No.:	1.49e-187	Length:	1824
Score:	2188.00	Matches:	416
Percent Similarity:	99.52%	Conservative:	1
Best Local Similarity:	99.28%	Mismatches:	2
Query Match:	99.18%	Indels:	0
DB:	6	Gaps:	0

US-10-798-532-4 (1-419) x AX680136 (1-1824)

QY 1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGlnAlaPheGly 20


```

Db      169 ATGGCGGAGATCCCGCTGCGCGCTGGGCGCGAGACCGCGCGCTTTCGAGGCGCTTCGCG
Qy      21 ArgLeuTTrGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
Db      229 CGGCTGTGGAGGAGTACAGACCGCTCTGGGTAAGCGCTCTCCCGCTCGAGTATCGGGTT 288
Qy      41 ArgCySeGlyAnPProGlySerProProGlyValAlaLeuArgGlnPheLeuProProGly 60
Db      289 CGTGTGGCGGCAACCTGGCTCGCCGCCCGCGCGCTCAAGAGATTCTTGGCGGCAAGA 348
Qy      61 ThrThGlyValAlaAlaSerAlaAlaGlyTyrGlyPheArgGlySerGlyAlaAlaLeu 80
Db      349 ACCACCGGGGCTCGCGCTCTGCGCGAGATATGTTTCCGCAAGAGAGCGGCGCTG 408
Qy      81 GlnGlnLeuGlnGlnHisArgAniLeuValThrLeuTyrGlyValPheThrIleHisPhe 100
Db      409 GAAACAGTTGACGGGTCAAGAAACATGTAATTTGTAAGAGTGTTCACATCCACTTT 468
Qy      101 SerProAnValProSerArgCySeLeuLeuLeuGlnLeuLeuAlaSerValSerGln 120
Db      469 TCTCCAAATGTGCATCAGCTGCTGTGCTTGAACCTCTGGATGTCAAGTGTTCGAA 528
Qy      121 LeuLeuLeuTyrSerSerHisGlnGlnGlyCySerSerMetTrpMetIleGlnHisGlyAlaArg 140
Db      529 TTGCTCTTATATTCAGGTCAACAGGGTGTTCATGTGATGATACAGCATTTGTGCCGA 588
Qy      141 AspValLeuGlnValAlaLeuValPheLeuHisHisGlnGlyTyrValHisAlaAspLeuVal 160
Db      589 GATGTTTGGAGGCCCTTCTCTTTCTTTCATCATGAGGCTATGTCATGCGAGCCTCMAA 648
Qy      161 ProArgAniLeuLeuTrpSerAlaGlnValGlnGlyCyAspPheLeuIleAspPheGlyLeu 180
Db      649 CCAAGTAACATATTTGGAGTGCAGAGAAATGAATGTTTAACTCATATGCACTTGGACTT 708
Qy      181 SerPheLeuGlnGlnValGlnValAspValGlyTyrIleGlnHisArgGlyTyrArgAlaPro 200
Db      709 AGCTTCAAAAGAGCAATCAGATGTAAATATATTCAGCAACACGGGATTCGGGCTCCA 768
Qy      201 GlnAlaGlnLeuGlnHisGlyAlaGlnAlaGlnGlyLeuGlnSerAspThrGlnGlyThr 220
Db      769 GAAGCGAATTCGAAATTCCTTGGCCGAGCTGGCTGCAAGTCAATCAAGATGACC 828
Qy      221 SerAlaValAspLeuTrpSerLeuGlnGlyIleIleLeuLeuGlnLeuTrpSerGlyMetLeu 240
Db      829 TCAGCTGTGATCTGTGAGACCTAGGAATCATTTTACTGGAATGTCTCAGGAATGAAA 888
Qy      241 LeuLeuHisThrValArgSerGlnGlnTrrPlyAlaAsnSerSerAlaIleIleAspHis 260
Db      889 CTGAAACATACAGTCAAGATCTCAGGAATGGAAGCAACAGTCTGCTATTAATGATCAC 948
Qy      261 IlePheAlaSerIleValAlaValAlaAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
Db      949 ATATTTCGCCGTAAGACAGTGTGATGCCGCAATTCACGCTTACCTTAAGACACTT 1008
Qy      281 IleLeuSerMetLeuHisAspAspProSerArgArgIleProAlaGlnLeuAlaLeuCyS 300
Db      1009 ATCAAAAGCATGCTTCATGATGATCCAAAGCAAGAAATTCCTGCTGAATGGCATTTGTGC 1068
Qy      301 SerProPhePheSerIleProPheAlaProHisIleGlnAspLeuValMetLeuProThr 320
Db      1069 AGCCCATTTCTTATGCAATTCCTTTTGGCCCTCATATGMAATGTGGCAAGCTTCCCACT 1128
Qy      321 ProValLeuArgLeuLeuAnValLeuAspAspAspTyrLeuGlnGlnGlnGlnGlyTyr 340
Db      1129 CCGTGTCTAAGATCTGTAATGTGCTGATGATGATTTATTTGAGATGAGAGGAAATAT 1188
Qy      341 GlnAspValValAlaGlnAspValGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 360
Db      1189 GAGGATGTTGTAGAAAGTGTAAAGAGGAGTGTCAAAATATATGGAACAGTGTATCTTA 1248
Qy      361 LeuValProLeuGlnAspProGlyValGlnGlnValPheValGlnTyrAlaAlaAlaGly 380

```

```

Db      1249 CTTGTTCCAAAGGAATCTTGACAGAGACAGATCTTGTGAGTATGCAAACTGCT 1308
Qy      381 AspSerIleValAlaGlnValLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
Db      1309 GATTCAAAGCTGCCGAGAAATTAATGACTGGAAGATGTTTGAATGGGAAGTTTGTG 1368
Qy      401 AlaThrPheTyrProLeuSerAlaTyrIleValGlnGlyTyrLeuTyrGlnThrLeu 419
Db      1369 GCTACATCTTACCGCTGAGTGTCTTACAAAGAGGAGATATCTGTATCAAACTTGTCTT 1425

RESULT 9
RANKISRNA
LOCUS      RANKISRNA      1633 bp      mRNA      linear      ROD 08-JUL-2002
DEFINITION R. norvegicus mRNA for KIS protein.
ACCESSION  X98374
VERSION    X98374.1 GI:1403531
KEYWORDS   KIS gene.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.

REFERENCE
1
AUTHORS    Maucuer, A., Ozon, S., Manceau, V., Gavet, O., Lawler, S., Curmi, P. and
            Sobel, A.
TITLE      KIS is a protein kinase with an RNA recognition motif
JOURNAL    J. Biol. Chem. 272 (37), 23151-23156 (1997)
MEDLINE    97435279
PUBMED     9287318
REFERENCE   2 (bases 1 to 1633)
AUTHORS    Maucuer, A.
TITLE      Direct Subdivision
SUBMITTED  (10-JUN-1996) A. Maucuer, INSERM U440, 17 rue du Fer
JOURNAL    Moulins, 75005 Paris, FRANCE
FEATURES
     source
         1..1633
             /organism="Rattus norvegicus"
             /mol_type="mRNA"
             /db_xref="taxon:10116"
             /tissue_type="brain"
             /clone_lib="7"
             /dev_stage="neonate"
             1..1633
             /gene="KIS"
             /gene="KIS"
             /gene="KIS"
             /gene="KIS"
             /codon_start=1
             /protein_id="CAA67021.1"
             /db_xref="GI:1403532"
             /db_xref="SwIsr-Pro:063285"
             /trna_allocation="MAGSGCAGAPPPFLFENFGRLMQVQSRLSGSSASVYRVCCG
             /TPGSPGALKQFLPGTTGGAASAEVYFRERALEQLOGBRNTVTLYGVETIHESP
             NVPSRCILIELIDVSVSLTLYSSHQGSMKMOICADYVLAFLHBBGVVADLK
             PNITLMSABECFKLIDFGLSFKSGNQVKITQTGTAPAEHLONCLAQAGSDTE
             CTSADVLSLGIILEMFSGMKLKRTVDSQEKWANSALIDHIFSKAVVNAALPAYH
             LRDLTKSMIHDDPSRRIRIPEAMALCSPFSSIPAPRIEDLVMLPTVVLNLVLDLDDYL
             ENDEDYEDVEDVEKECOKYGVVSLVLPKPNRGQYFVEVYANAGDSKAQKLLTGR
             MPDKFVATFPLPSAYKRGVLYQTLL"

ORIGIN
Alignment Scores:
Pred. No.:      3,61e-187      Length:      1633
Score:          2183.00      Matches:      414
Percent Similarity: 99.52%      Conservative: 3
Best Local Similarity: 98.81%      Mismatches:  2
Query Match:    98.96%      Indels:      0
DB:             10      Gaps:      0

US-10-798-532-4 (1-419) x RANKISRNA (1-1633)
Qy      1 MetAlaGlySerGlyCyAlaAlaTrpGlyAlaGluProProArgPheLeuGlnAlaPheGly 20

```



```

Db      39 ATGGCCGGGTCGGCTCGCGTGGGGCGCCAGCCCGCGCTTCTGAGGCTTCGGG 98
Qy      21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
Db      99 CGGCTGTGGAGGTACAGACCGCGCTTAGCGAGGGCTCTCGGCGTGGTGTCCGGGTG 158
Qy      41 ArgCysCysGlyValanProGlySerProProGlyValAlaLeuArgLysPheLeuProGly 60
Db      159 CGGTGCTGGCGGCACTCCAGGCTGGCCCCCGGCGGCGCTCAAGCATTCCTCGCTCCGGGA 218
Qy      61 ThrThrValAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
Db      219 ACCACCGGGGCTCGCGCTCGCGCTGCGAGTATGTCTTCGAAAGAGAGAGCGAGCGCTG 278
Qy      81 GluGluLeuGlnGlnLysAspGlnLeuValThrLeuTyrGlyValAlaPheThrLysPhe 100
Db      279 GAGCAGTGGAGGGTCAAGAGAACATCGTACCTTTATACGAGATCTTTACCATACACTTC 338
Qy      101 SerProAsnValProSerArgCysLeuLeuGluLeuLeuAspValSerValSerGlu 120
Db      339 TCTCCAAATGTGCATCAACGCTGTCTGTCTTGAAGTCTCGATGTGCAAGCTTCGGAA 398
Qy      121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetLysGlnHisCysAlaArg 140
Db      399 TTGCTGTATATATCCAGTCAACAGGGTTCCTCATGTGATGATCCAGCACTGTGCCAGA 458
Qy      141 AspValLeuGlnValAlaLeuAlaPheLeuHisGlnGlyTyrTyrAlaHisAlaAspLeu 160
Db      459 GATGCTCTGGAGGCCCTTGTCTTTCTTTCACCATGAGGGCTATGTCCATGAGACCTCAAG 518
Qy      161 ProArgAsnIleLeuTrpSerAlaGluValAsnGluCysPheLysLeuIleAspPheGlyLeu 180
Db      519 CCAGCAAAACCTCTGTGAGTGGCGAGAACGAGTCTTTAAAGCTTATATGACTTTGAGATC 578
Qy      181 SerPheLysGlnGlyAsnGlnAspValLysTyrLysGlnThrAspGlyTyrArgAlaPro 200
Db      579 AGCTTCAAAAGAGGCAATCAGAGCGTGAAGTATATTCAGACAGCGGGTATCCAGCTCT 638
Qy      201 GluAlaGluLeuGlnAsnGlyLeuAlaGlnAlaGlyLeuGlnLeuAspThrGluCysThr 220
Db      639 GAAGCAAAATGAGATTCCTTGTCCAGCGCGGCTGAGAGTGTACAGAGTGTACC 698
Qy      221 SerAlaValAspLeuTrpSerLeuGlyLysIleIleLeuLeuGlnMetPheSerGlyMetLys 240
Db      699 TCGAGCTTGTATCTCTGAGGCTTGGAAATCTTTACTGAAATTTCTTCAGAAATGAAA 758
Qy      241 LeuLysHisThrValArgSerGlnGluTyrLysAlaAsnSerSerAlaIleIleAspHis 260
Db      759 CTGAAACATACAGTCAAGATCTCAGAGAGTGAAGCAACAGTTCTGCTATTTATGATCAT 818
Qy      261 IlePheAlaSerLysAlaValAlaAspAlaAlaIleProAlaTyrHisLeuAlaGlyLeu 280
Db      819 ATATTTGGCCAGTAAAGAGTGGTGAAGCGCAATTCACACCTATCACTCAAGAGACTT 878
Qy      281 IleLysSerMetLeuHisAspAspProSerArgArgLysProAlaGluMetAlaLeuCys 300
Db      879 ATCAAAAGCATGCTTCATGACGACCCAGCAAGAAAGATCCCTGCTGAGATGGCTTGTGC 938
Qy      301 SerProPhePheSerLysProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
Db      939 AGCCCAATTTCTTAGCATTTCTTTTGGCCCTCATATGAAATGATCGTGTCCAACT 998
Qy      321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluTyr 340
Db      999 CGAGTGTCAAGCTCCCAATGTGCGAGATGATGATATCTTGAATGAAGATGATAT 1058
Qy      341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
Db      1059 GAAGATTTTGTAAAGATGTAAAGAGAGATGTCAAGAAATGTGACCAAGTGTCTCTG 1118
Qy      361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
Db      1119 CTGTGTCAAAGAAATCTCTGGCAGAGCAAGTCTTTGTGAGTACGCAAAAGCTGGA 1178

```

```

Qy      381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
Db      1179 GATTCAAAGCTGCTCAAGAAATTCCTGACTGGAGAGATTTTGAACGAAATTTGTGTG 1238
Qy      401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
Db      1239 GCTACATTTACCCGTGAGTGTGCTTACAGAGGGATATCTTTATCAAACTTGCTT 1295

RESULT 10
RNU70372
LOCUS
DEFINITION
Rattus norvegicus PAM COOH-terminal interactor protein 2 mRNA,
complete cds.
VERSION
U70372.2 GI:5821767
KEYWORDS
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
TITLE
1 (bases 1 to 3244)
Alam,M.R., Caldwell,B.D., Johnson,R.C., Darlington,D.N., Mains,R.E.
and Bipper,B.A.
The Novel Kinase P-CIP2 Interacts with the Cytosolic Routing
Determinants of the Peptide Processing Enzyme Peptidylglycine
alpha-Amidating Monooxygenase
J. Biol. Chem. 271 (1996) In press
3 (bases 1 to 3244)
Alam,R., Caldwell,B.D., Johnson,R.C., Darlington,D.N., Mains,R.E.
and Bipper,B.A.
Direct Substitution
Submitted (11-SEP-1996) Neuroscience, Johns Hopkins University
School of Medicine, 725 N. Wolfe Street, WBSB 902A, Baltimore, MD
21205, USA
4 (bases 1 to 3244)
Caldwell,B.D., Darlington,D.N., Penzes,P., Johnson,R.C.,
Bipper,B.A. and Mains,R.E.
Direct Substitution
Submitted (03-SEP-1999) Neuroscience, Johns Hopkins University
School of Medicine, 725 N. Wolfe Street, WBSB 902A, Baltimore, MD
21205, USA
REMARK
Sequence update by submitter
On Sep 3, 1999 this sequence version replaced gi:1698778.
FEATURES
source
CDS
/mol_type="mRNA"
/db_xref="taxon:10116"
162..1421
/function="serine protein kinase"
/notes="P-CIP2"
/codon_start=1
/product="PAM COOH-terminal interactor protein 2"
/protein_id="AAC53031.2"
/db_xref="GI:5821768"
/translation="MAGSCAWGASGAPPFLAFLRLMQVQSLGSSASVYRVACCG
TPGSPGALKPLPGTTGAASAAEYGRKERALBOLQGRNVTLYGVTTIFSP
NYPSCILIEFDIVSSELYVSSHQGSMWIOHCAVDVLEALFLHREGVHADIK
PNITLKSANECFKLIDFGLSKFEGNDVKYTIQTGCTRAPAELEONCAQAQSGDTE
CTSAVDLWSLGIILEPFGMKLKHIVSQEKANSSAIIHIFASKVAVNNAIPIYH
LNDLIKSMHDDPSRRIPAEMLACSPFSIIPAPRIEDLVMLPTVRLANVLDDYL
ENDEYEDVDEVEKEQCYGIVVSLVLPKBNPBGQYFVEYANAGDSKAAQKLTGR

```

ORIGIN

MFDGKRVATFYPPLSAKRGVLYQTLL"

Alignment Scores:

Pred. No.:	9,42e-187	Length:	3244
Score:	2183.00	Matches:	414
Percent Similarity:	99.52%	Conservative:	3
Best Local Similarity:	98.81%	Mismatches:	2
Query Match:	98.96%	Indels:	0
DB:	10	Gaps:	0

US-10-798-532-4 (1-419) x RNUT0372 (1-3244)

QY 1 MetAlaGlySerGlyCysAlaATPGLyAlaGluProProArgPheLeuGluAlaPheGly 20
 DB 162 ATGGCGGGGCTCGGGCTCGGGGGGCGGAGCCGCGGCTTCCGAGGGCTTCGGG 221
 QY 21 ArgLeuTPGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyraGVal 40
 DB 222 CGGCTGGGAGGATACAGAGCCGCTTAGGCGAGCGGCTTCGGGCTCGGGTACCGGGTG 281
 QY 41 ArgCysCysGlyAAsnProGlySerProProGlyValAlaLeuArgLInPheLeuProProGly 60
 DB 282 CGCTGCTGCGGCACTCCAGGCTCGCCCCCGGCGCCCTCAAGCAGTTCCTGCTCCGGA 341
 QY 61 ThrThyGlyAlaAlaAlaSerAlaAlaGluTyrglyPheArgGlyGluArgAlaAlaLeu 80
 DB 342 ACCACCGGGGCTCGGCTCGGGTGGAGTATGTTCCGAAAGAGAGCGAGCGCTG 401
 QY 81 GluGlnLeuGlnGlyValGatGanLLeValThrLeuTyrglyValPheThrLLehPhe 100
 DB 402 GAGCAGATGGAGGATCAGAGAACATGCTACTTTATACGAGATCTTTACCATACACTTC 461
 QY 101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
 DB 462 TCTCCAAATGTGCCATACGCTGCTGTTCTTGAACCTCGAGATGTCAGCGTTGGAA 521
 QY 121 LeuLeuLeuTySerSerArgGlnGlyCysSerMetTrpMetLLeGlnHisCysAlaArg 140
 DB 522 TTGCTGTATATTCAGTCCAGGCTCGGGTGTCTTCATGTGATGATCCAGCATGTCGCA 581
 QY 141 AspValLeuGluAlaLeuAlaPheLeuHisGlnGlyTyTyraHisAlaAspLeuVal 160
 DB 582 GATGCTCTGAGGCGCTTGTCTTCTTCTTCAACATGAGGGCTATGTCATGACAGCTTCAG 641
 QY 161 ProArgAsnLLeuLeuTPSerAlaGluAAsnGlyCysPheLeuLeuLeuPheGlyLeu 180
 DB 642 CCAAGAAACATCTCTGAGAGTGGGAGACGAGTGTCTTAAGCTTATGACTTGGACTC 701
 QY 181 SerPheGlyGlnGlyAAsnGlnAspValTyTyraLLeGlnThrAspGlyTyraGAlaPro 200
 DB 702 AGCTTCAAGAGAGCAATCAGAGAGCTGATATATTCAGACAGCGGATATCAGACTCCT 761
 QY 201 GluAlaGlnLeuGlnAAsnGlyLeuAlaGlnAlaGlyLeuGlnSerAspThrGlyCysThr 220
 DB 762 GAGCGAGAACTGAGATTCCTTGCCCGGAGCGGCTTGAGAGTATACAGAGTACC 821
 QY 221 SerAlaValAspLeuTPSerLeuGlyLLeLLeuLeuGlnLeuPheSerGlyMetLeu 240
 DB 822 TCAGCTGTATCTCTGAGACTTGGAAATCATTTTACTGAAATGTTCTTCAAGAAAGAAA 881
 QY 241 LeuLysHisThrValArgSerGlnGluTyrglyValAlaAsnSerSerAlaLLeLLeuPHis 260
 DB 882 CTGAACATATCAGATCAGATCTCAGAGATGGAAGCAACAGTTCTGCTATATTAATCAT 941
 QY 261 LLePheAlaSerTyAlaValAlaValAlaAlaAlaLLeProAlaTyHisLLeuAlaGAspLeu 280
 DB 942 ATATTGCGCGATGAAGAGTGTGAATGCGCAATTCACGCTTACCTCAAGAGACTT 1001
 QY 281 LLeLysSerMetLeuHisAspAspProSerArgArgLLeProAlaGluMetAlaLeuCys 300
 DB 1002 ATCAAAAGCATGCTTCTCATGAGCAACCAAGCAAGAGATCCCTGCTGAGATGCGCTTGTC 1061

QY 301 SerProPheSerLLeProPheAlaProHisLLeGluAspLeuValMetLeuProThr 320
 DB 1062 AGCCCATCTTATGATCTCTTTCCTCCCTCATATTAAGATCTGATGCTTCCAACT 1121
 QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspArgTyrglyValAsnGlnGlnGly 340
 DB 1122 CCAGTCTAGACTCTCCATATGCTGATGATGATCTTATGAAATGAAATGAAATAT 1181
 QY 341 GluAspValAlaGlnAspValLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 360
 DB 1182 GAAAGATCTTGAAAGATTAAGAGAGCTCAGAAATATGACACAGTGTCTTCTCG 1241
 QY 361 LeuValProGlyGluAsnProGlyValArgGlyGlnValPheValGluTyraAlaAsnAlaGly 380
 DB 1242 CTGTGTCGAAAGAAATCTGCGAGAGCAAGCTTGTGTGATGACGCAAAACCTGGA 1301
 QY 381 AspSerValAlaGlnLysLeuLeuThrGlyValGlyMetPheAspGlyLysPheValVal 400
 DB 1302 GATTCGAAAGCTGCTCAGAAATGCTGACTGGAGAGATTTGACGGAGAGTTTGTG 1361
 QY 401 AlaThrPheTyProLeuSerAlaTyrglyArgGlyTyrglyLeuTyrglnThrLeuLeu 419
 DB 1362 GCTACATTTTACCGCTGAGTGCCTTACAGAGGGATATCTTATCAAACTTGCTT 1418

RESULT 11

BC058732 2041 bp mRNA linear ROD 30-JUN-2004
 LOCUS BC058732.1 GI:37194892
 DEFINITION Mus musculus kinase interacting with leukemia-associated gene (setd4h1n), mRNA (cDNA clone MGC:64770 IMAGE:6414877), complete cds.

ACCESSION BC058732
 VERSION BC058732.1
 KEYWORDS GC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE

1 (bases 1 to 2041)
 Strausberg, R.D., Feingold, E.A., Grouse, L.H., Derge, J.G., Altschul, S.F., Zeeberg, B., Buetow, K.H., Scheffer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stappleton, M., Soares, M.B., Bonaldo, A.F., Cavaletto, T.L., Scheetz, T.E., Brownstein, M.J., Ueda, T.B., Toshimune, S., Carninci, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullah, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Schevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E., Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

TITLE

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITILE

JOURNAL

REMARK

COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLB)

DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant.

Web site: <http://genome.uiowa.edu>

Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
Bonaldo, M.F., Akabogu, I., Bair, T., Bair, J., Crouch, K., Davis, A., Fishler, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K., Schetz, T., Smith, C., Snit, E., Tack, D., Trout, K., Walters, J., Casavant, T., Soares, M.B.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLN at: <http://image.llnl.gov>
Series: Plate: Row: Column: 0
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6754447.

FEATURES

source

1. 2041
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="MGC:64770 IMAGE:6414877"
/tissue_type="Brain, pool of enriched mouse brain tissues from 13.5, 14.5, 16.5 and 17.5 dpc"
/clone_id="N1H BMAP_FRO"
/lab_host="DH10B"
/note="Vector: pYX-ASC"

gene

1. 2041
/gene="Kist"
/note="synonym: KIS"
/db_xref="LocusID:16589"
/db_xref="MGI:1341908"
154. 1413
/gene="Kist"
/product="Kinasase interacting with leukemia-associated gene (statmin)"
/protein_id="AAH58732.1"
/db_xref="GI:37194893"
/db_xref="LocusID:16589"
/db_xref="MGI:1341908"

CDS

/codon_start=1
/product="Kinasase interacting with leukemia-associated gene (statmin)"
/protein_id="AAH58732.1"
/db_xref="GI:37194893"
/db_xref="LocusID:16589"
/db_xref="MGI:1341908"
/translaction="MAGSGCAGWABPRFLAFLGMVQSRIGSSSASVYKRVCCGTPGSPGALQKFLPGTGTGAAAGAAVYGRKERALEQLQGHNIYTLGVFTIFSPNVSRCLELLELVSEBLLYSHOGSCMMIOHOCARVLEALFLHGGVYADLKPNILWASNEBECKLIDPLSEKGNODVXYIOTDGRVAPRELQCLADAGISDTECSAYDNLKSIGITLLEMFSGMKKHTYRSGEMKANSATIDHIFAKAYVNAIPIYKHLRLKSMLEHDDPKRIIPALMALCSPFSPFPAVHLEDVMLPTVLRLLNVDDLYLEDEYEDVVEDYKECCOKRGVPSVLLVPENRGQVVEYVANAQDSKAAQKLTLGRMEDGKEFVATFYPLSAVYKRGYLTLL"

ORIGIN

Alignment Scores:

Pred. No.: 6.07e-187 Length: 2041
Score: 2182.00 Matches: 414
Percent Similarity: 99.28% Conservative: 2
Best Local Similarity: 98.81% Mismatches: 3
Query Match: 98.91% Gaps: 0
DB: 10

US-10-798-532-4 (1-419) x BC058732 (1-2041)

QY 1 MetAAG1SerG1yCysa1aTpg1yAlaGluProProArghPheLeuGluAlaPheGly 20
DB 154 ATGGCGGGGTCGGGCTGCGGCGGGGCGGAGCCGCGGCTTCTTGAAGGCTTCGGG 213
QY 21 ArgLeuTpg1nValGlnSerArGleuG1ySerG1ySerSerAlaSerValTyArGval 40
DB 214 CGGCTGAGGAGGTCCAGAGCCGCTGGGCGAGCGGCTCGTCCGCTCGGTACCGGGTG 273
QY 41 ArgCyeGeg1yAnPrG1ySerProProG1yAlaLeuArgGlnPheLeuProProGly 60
DB 274 CGGTGCTGGCGGTACCCGCGGCTCGCCCGCGGCGCTCAAGAGTTCTCGCTCGGGGA 333
QY 61 ThrThrGlyAlaAlaSerAlaAlaGluTyGlyPheArgG1yGluArgAlaAlaLeu 80

DB 334 ACCACGGGGCTGCGGCTCGGCGGAGTATGTTCCGCAAGAGGGCGGCTG 393
QY 81 GluGlnLeuGlnGlyN1hArgAnT1LeValThrLeuTyGlyValPheThr1Le1h1Phe 100
DB 394 GAGCAGTTGCAGGGTCACAGGAACATCGACTTATATAGAGTCTTATCAACATTC 453
QY 101 SerProAnValProSerArgCysLeuLeuGluGluLeuAspValSerValSerGlu 120
DB 454 TCTCCCAATGTCATCAGCTGTCTGTCTTGAATCTTGATGTCACTGTTCGGA 513
QY 121 LeuLeuLeuTySerSerN1hGlnGlyCysSerMetTrpMet1LeGlnN1hCysAlaArg 140
DB 514 TTGCTTTATATTCAGTCATCAGGGCTGCTCATGTGATGATACAGACTGTGCCGA 573
QY 141 AspValLeuGluAlaLeuAlaPheLeuN1h1hGlnGlyTyTyValN1hAlaAspLeuTyS 160
DB 574 GATGTTCTGAGGCCCTCGCTTCTTTCACCATGAGGGCTACGTCCATGACACCTCAA 633
QY 161 ProArgAn1LeuTTrpSerAlaGluAnGluCysPheTyLeu1LeAspPheGlyLeu 180
DB 634 CACGAAACATCTGTGAGTCCGAGAAATGAGTGTTTTAACTTATGACTTTGACTC 693
QY 181 SerPheTyG1yG1yAnG1nAspValTyTy1LeGlnThrAspG1yTyTyArgAlaPro 200
DB 694 AGCTTCAAGAAAGGCAATCAGAGACGTAAAGTATATTCAGACAGAGGATGCGCTCCT 753
QY 201 GluAlaGluLeuGlnAnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGlyTyThr 220
DB 754 GAAGCAGAGCTCAGAACTGCTTGGCCCGAGCGGCTGACAGAGTACAAAGTAC 813
QY 221 SerAlaValAspLeuTTrpSerGly1Le1LeuLeuGluMetPheSerG1yMetTyS 240
DB 814 TCAAGCTGATGTGAGAGCTCGAATCATTTACGAAATGTTCTCAGAAATGAAG 873
QY 241 LeuTyN1h1eThValArgSerGlnGlyTyTyValAsnSerSerAla1Le1LeAspN1h 260
DB 874 CTGAACACATACAGTCAGATCTCAGAGATGGAAGCAAGTTCGCTATTTATGATCAT 933
QY 261 1LePheAlaSerTyAlaValAlaValAlaAla1LeProAlaTyTh1eLeuArgAspLeu 280
DB 934 ATATTGTCAGATGAAGCAGATGTGATGCGCAATTCAGCTTACCTCAGAGACTT 993
QY 281 1LeTySerMetLeuN1hAspAspProSerAlaGlu1LeProAlaGluMetAlaLeuTyS 300
DB 994 ATCAAAAGATCTTCATGATGATACCCGGGCAAGAAATCCCTGCTGAGATGSCATTGTGC 1053
QY 301 SerProPhePheSer1LeProPheAlaProN1h1LeGluAspLeuValMetLeuProThr 320
DB 1054 AGCCCATTTCTTATAGCATTCCTTTGCCCCCTCATATTGAAGATCTGTGATGCTCGACT 1113
QY 321 ProValLeuArgLeuLeuAnValLeuAspAspAspTyTyLeuG1yAnG1yGluTyTy 340
DB 1114 CCAGTGCTCAGGCTCCTCATGTGCTGATGATGATTACTTGAAGAAATGAATGAATAT 1173
QY 341 GluAspValValGluAspValTyTyGluTyTyGlnTyTyGlyTyProValValSerLeu 360
DB 1174 GAAGATGTTGAGAGATGTGAAGAGAGGTCTCAAGAAATATGACACAGTGTCTCTCG 1233
QY 361 LeuValProTyG1yAnPrG1yArgG1yGlnValPheValGluTyTyAlaAnAlaGly 380
DB 1234 CTTGTTCCAAAGAAATCTTGCAAGAGCAAGTCTTGATGATACCGCAACCTGCTGT 1293
QY 381 AspSerTyAlaAlaGlnTyLeuLeuTyTyArgMetPheAspG1yTyPheValVal 400
DB 1294 GATTCCAAAGCTGCTCAGAAATGCTGACTGGAGAGATGTTTGAAGGAAAGTTTGTGTG 1353
QY 401 AlaThrPheTyTyProLeuSerAlaTyTyTyAspG1yTyTyLeuTyTyGlnThrLeuLeu 419
DB 1354 GCTACATTCATCCCGCTGAGTGTCTCAAGAGGAGATATCTTTATCAAACTTGCTT 1410
RESULT 12
AC129141_1/c

WPCOMMENT
 Sequence split into 4 fragments LOCUS AC129141 Accession AC129141
 Fragment Name Begin End
 AC129141_0 1 110000
 AC129141_1 100001 210000
 AC129141_2 200001 310000
 AC129141_3 300001 363206
 Continuation (2 of 4) of AC129141 from base 100001 (AC129141 Rattus norvegicus clone CH2)

Alignment Scores:
 Prod. No.: 8 44e-184 Length: 110000
 Score: 2174.00 Matches: 413
 Percent Similarity: 99.28% Conservative: 3
 Best Local Similarity: 98.57% Mismatches: 3
 Query Match: 98.55% Indels: 0
 DB: 2 Gaps: 0

US-10-798-532-4 (1-419) x AC129141_1 (1-110000)

QY	1	MetAlaGlySerGlyCysAlaATPglValaGluProPheArgPheLeuGluAlaPheGly	20
DB	70316	ATGAGCGGGGCTGCTGCTGCGGCGGGGCGGAGCGCGCGCTTCTGAGAGCCTTCGGG	70257
QY	21	ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal	40
DB	70256	CGGCTGGGAGGTACAGAGCCGCTAGGAGCGGCTCTCGGCTCGGTGACCGGGTG	70197
QY	41	ArgCysGlyValAsnProGlySerProProGlyValAlaLeuArgPheLeuProProGly	60
DB	70196	CGCTGCTGGGCGACTCCAGGCTGCGCCCGCGGCGCTTCAAGAGTTCCTCCGCGGA	70137
QY	61	ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgGlyValArgAlaLeu	80
DB	70136	ACCAAGCGGGGCTGCGGCTGCGGCTGCGGAGTATGTTTCGAAAGAGAGGAGCGCTG	70077
QY	81	GluGluLeuGlnGlnHisArgAsnIleValThrLeuTyrGlyValaPheTrpIleHisPhe	100
DB	70076	GAGCAGTGGAGGATCAGAGAACTGTAATTAACGAGCTTTACATACATACACTTC	70017
QY	101	SerProAsnValProSerArgCysLeuLeuGluGluLeuAspValSerValSerGlu	120
DB	70016	TCTCCAAATGTCATACACCTGCTGCTTCTTGAACCTCGATGTCAGCCTTCGGA	69957
QY	121	LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg	140
DB	69956	TTGCTGTATATATCCAGTCAACGAGGTGCTCATGTGATGATACGACACTGCGAGA	69897
QY	141	AspValLeuGluAlaLeuAlaPheLeuHisGluGlyTyrValHisAlaAspLeuVal	160
DB	69896	GATGTTCTGGAGGCTTGTCTTTTTCACCATGAGGGCTATGTCATGCACTCAAA	69837
QY	161	ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLeuLeuAspPheGlyLeu	180
DB	69836	CCAGCAAACTCTCTGAGAGTGGAGAAAGAGTCTTTAAGCTTATGACTTGAATC	69777
QY	181	SerPheLeuGluGlnValAsnValValTyrTyrIleGlnThrAspGlyTyrArgAlaPro	200
DB	69776	AGCTTCAAGAGAGCAATCAGAGCTGAGTAATATTCACAAACGGGTATTCAGACTC	69717
QY	201	GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspTrpGluCysThr	220
DB	69716	GAGGAGAACTGCAAGATTCCTTGCCGAGGCGGCTGCAAGATGATACAGAGTACC	69657
QY	221	SerAlaValaAspLeuTrpSerLeuGlyIleIleLeuLeuGluLeuPheSerGlyMetVal	240
DB	69656	TCAAGCTTATCTCTGAGACCTTGGAAATCATTTTACTGAAATGTTCTCAGAAATGAA	69597
QY	241	LeuLeuHisThrValaArgSerGlnGluTyrPylValaAsnSerSerAlaIleIleAspHis	260
DB	69596	CTGAAACATACAGTCAAGTCTCAGAGTGAAGCAAAAGTTCTGCTATATATGATCAT	69537
QY	261	IlePheAlaSerValaValaValaAsnAlaIleProAlaTyrHisLeuArgAspLeu	280

DB	69536	ATATTTGCCAGTAAGACAGTGTGAATGCCGCAATTCACGCTTACCTCAGACACTT	69477
QY	281	IleYsSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys	300
DB	69476	ATCAAAAGCATCTTCATATGACACCCAGACAGAGATCCCTGCTGAGATGCGCTTGCC	69417
QY	301	SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr	320
DB	69416	AGCCCATCTTTAGATTCCTTTTGCCCTCATATATGAGATCTGTGATGCTTCAACT	69357
QY	321	ProValLeuArgLeuLeuAsnValLeuAspAspTyrLeuGlyValAsnGluGluTyr	340
DB	69356	CAAGTCTAGACTCTCATATGCTGATGATGACTATCTTGAATACAAATGAATATAT	69297
QY	341	GluAspValaValaGluAspValaValaGluGluCysGlnValTyrGlyProValaValSer	360
DB	69296	GAAAGATGTTAGAAAGATGTAAAGAGAGTGTCAAAATATGACCAAGTGTCTCTCG	69237
QY	361	LeuValProValGluAsnProGlyValArgGlyGlnValPheValGluTyrAlaAsnAlaGly	380
DB	69236	CTTGTTCCAAAGAAATCTGCGAGAGCAAGTCTTGTGATGATGCGCAACCGCTGGA	69177
QY	381	AspSerValaAlaGlnValLeuLeuThrGlyArgMetPheAspGlyLysPheValaVal	400
DB	69176	GATTCCAAAGCTGCTCAGAAATGCTGACTGAGAGATGTTGATGGAAAGTTGTGTG	69117
QY	401	AlaThrPheTyrProLeuSerAlaTyrIlyAspArgGlyTyrLeuTyrGlnThrLeuLeu	419
DB	69116	GCTACATTTACCGGTGATGCTCTACAGAGAGGATATCTTATCAAACTTGTCT	69060

RESULT 13
 MPEKIS
 LOCUS
 DEFINITION
 accession
 Y10725
 Y10725.2 GI:21726712
 keywords
 KIS gene; protein kinase.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1254)
 Maucuer, A., Ozon, S., Manceau, V., Gavet, O., Lawler, S., Curmi, P. and
 Sobel, A.
 KIS is a protein kinase with an RNA recognition motif
 J. Biol. Chem. 272 (37), 23151-23156 (1997)
 MEDLINE
 PUBMED
 9287318
 REFERENCE
 2
 Maucuer, A.
 Direct Submision
 Submitted (23-JAN-1997) Maucuer A., INSERM U440, U440, 17 Rue du
 Fer Moulin, Paris, 75005, FRANCE
 JOURNAL
 revised by [4]
 REMARK
 3 (bases 1 to 1254)
 Maucuer, A.
 Direct Submision
 Submitted (02-JUL-2002) A. Maucuer, INSERM U440, U440, 17 Rue du
 Fer Moulin, Paris, 75005, FRANCE
 COMMENT
 On Jul 10, 2002 this sequence version replaced gi:1806129.
 Overlaps with related sequences X82320, X10725.
 FEATURES
 source
 1..1254
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="K22"
 /issue_type="brain"
 /clone_id="lambda gcl10"
 /dev_stage="neonate"
 1..1254
 /gene="Kis"
 11..>1254
 CDS

```

/gene="Kis"
/codon_start=1
/product="Kis protein kinase"
/protein_id="CAA71714.2"
/db_xref="GI:21726713"
/db_xref="GGA:P97343"
/db_xref="Swiss-Prot:P97343"
/translation="MAGSGCAGMAGPRLFAARGLMVOYRSLGSGSSAYVRVRCGG
TRGSPRGALAKQPLPRGTGAAASAAAYGRKEAALFQAGHNITTYLIGVPTIHSP
NPFSRLLELDVSVSELDYSSHQGCSMMIIOHCARDVLEALATHHGIVHALDK
PRLIMASNECKFLIDFGLSFREGQDVYIOTDGRAPAEALQNCLOAGQSTPE
CTSAVDLMSLIGILLEMFGMKLHTVRSGEMKANSALIDHIFASKAVVNAIIPYH
LRLIKSMLEHDDPGRIPAEMLKSPFSLPFAPIHIEDLVMLPTPLRLINLVDDLYL
ENEDVEDVVEDYKRCOKYGPVSLIVPRENRRGQVREYVANNADSKAAQKLLTRGR
MDGRKRVVATFPLSAIKKGYL"

```

ORIGIN

Alignment Scores:

```

Pred. No.:      5,55e-185      Length:      1254
Score:          2157.00      Matches:      409
Percent Similarity: 99.28%      Conservative: 2
Best Local Similarity: 98.79%      Mismatches: 3
Query Match:     97.78%      Indels:      0
DB:              10      Gaps:      0

```

US-10-798-532-4 (1-419) x MMPKIS (1-1254)

```

QY      1 MetAlGlySerGlyCybAlaATPGlYAlaGluProProArGpHeuGluAlaPheGly 20
DB      11 ATGGCGGGGCTCCGCTGCGCGGGGCGCGAGCCGCGCTTCTCGAGGCGCTTCGGG 70
QY      21 ArgLeuTPGlnAlaGlnSerArgLeuGlySerGlySerSerAlaSerValYTrArgVal 40
DB      71 CGGCTGGGAGGTCAGAGCCGCTGGGCGAGGCGCTGCGGCTGCGGTACCGGCTG 130
QY      41 ArgCySeGlyAAspProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
DB      131 CGGTGCTGGGGTACCCGGGCTGCGCCCGCGGCGCTCAAGAGATTCCGCTCCGGGA 190
QY      61 ThrThGlyAlaAlaAlaSerAlaAlaGluTrGlyPheArgGlyGluArgAlaLeu 80
DB      191 ACCACCGGGGCTCGGCTCGCGCGCGAGTATGTTTCCGCAAGAGAGGCGGCGCTG 250
QY      81 GluGlnLeuGlnGlnHisArgAsnIleValThrLeuTrGlyValPheThrIleHsphe 100
DB      251 GAGCAGATTGAGGGGTACAGAGAAATCTGTACTTTATATGAATCTTTACCATACACTTC 310
QY      101 SerProAsnValProSerArgCybLeuLeuGluLeuLeuAspValSerValSerGly 120
DB      311 TCTCCCAATGTGCATACACGCTGTCTGCTTGAACCTCTGATGTCAGTCTTCGGAA 370
QY      121 LeuLeuLeuTrSerSerHisGlnGlyCybSerMetTrpMetIleGlnHisCybAlaArg 140
DB      371 TTGCTTTTATATCCAGTCAATCAGGCGTCTCCATGTGATGATACAGCACTGTGCCGA 430
QY      141 AspValLeuGlnAlaLeuAlaPheLeuHisGlnGlyTrpValHisAlaAspLeuVal 160
DB      431 GATGTTCTGAGGCGCTCGCTTTTTCACCAATGAGGCGTACGTCATGACAGACTCAAA 490
QY      161 ProArgAsnIleLeuTrpSerAlaGluAsnGlyCybPheLeuLeuLeuAspPheGlyLeu 180
DB      491 CCAGCAAAATCTCTGTGAGACTGCGAGAAATGATGTTTAAAGCTTATGACTTTGAGCTC 550
QY      181 SerPheLeuGlnGlyAsnGlnAspValYTrIleGlnThrAspGlyTrpArgAlaPro 200
DB      551 AGCTTAAAGAAAGGCATCAGACGTAATACTATATTCAGACAGCGGGTATCCCGCTCT 610
QY      201 GluAlaGluLeuGlnAsnGlyAlaGlnAlaGlyLeuLeuLeuAspThrGlnGlyThr 220
DB      611 GAAAGCAGAGCTGAGAACTCTTGGCCAGGCGGCTGAGAGTGAATACAGAAATGACC 670
QY      221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnLeuPheSerGlyMetLeu 240

```

```

DB      671 TCAGCTGTATGATCTGTGAGCCTCGGAATCATTTTACTGGAATGTCTCAGAAATGAG 730
QY      241 LeuLeuHisThrValArgSerGlnGluTrpLeuAlaAspSerSerAlaIleIleAspHis 260
DB      731 CTGAAACATACAGTCACTGATCTCAGGAATGAAAGGCAACAGTTCGTATTTATGATCAT 790
QY      261 IlePheAlaSerLeuAlaValAlaAsnAlaAlaIleProAlaTrpHisLeuArgAspLeu 280
DB      791 AATTTCGCAATGAACAGATGCTGAATGCCCAATTCAGCTTATCACTCAGAGACTT 850
QY      281 IleLeuSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCyb 300
DB      851 ATCAAAAGCATGCTTCATGATGACCCGGCAGAGAAATCCCTGCTGAGATGACATGTGCG 910
QY      301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
DB      911 AGCCCATTCCTTAAAGCATTCCTTTTGGCCCTCAATATGAAGATCTGTGTGCTTCGACT 970
QY      321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTrpLeuGlyAsnGlnGluTrpYr 340
DB      971 CCACTGCTCAAGGCTCTCTCATGTGCTGATGATGATGATACCTTGAAATGAAATGAAATAT 1030
QY      341 GluAspValValGluAspValIleGlnGluGlySerGlnLeuTrpGlyProValValSerLeu 360
DB      1031 GAAAGATGTTGTGAAGATGTAAGAGAGAGTGCAGAAATATGACCACTGATGTTCTCTG 1090
QY      361 LeuValProLeuGluAspProGlyArgGlnValPheValGluTrpAlaAsnAlaGly 380
DB      1091 CTGTTCCTCAAGAGAAATCTGCGCAGAGACAGCAAGCTTCTGTGAGTACCGCAACGTGAT 1150
QY      381 AspSerLeuAlaAlaGlnLeuLeuLeuThrGlyArgMetPheAspGlyLeuPheValVal 400
DB      1151 GATTCCAAGCTGCTCAAGAGTTCCTGACTGCGAGAGATGTTTGACCGGAAATGTTGTG 1210
QY      401 AlaThrPheTrpProLeuSerAlaTrpYrLeuArgGlyTrpLeu 414
DB      1211 GCTACATTCACCGCTGAGTGCCTACAGAGAGGATATCTT 1252

RESULT 14
BC026046      1367 bp      mRNA      linear      PRI 30-JUN-2004
LOCUS      BC026046.1      GI:19684094
DEFINITION      Homo sapiens kinase interacting with leukemia-associated gene
                  (leukhmn), mRNA (cDNA clone MGC:39391 IMAGE:4649261), complete
                  cds.
ACCESSION      EC026046
VERSION      BC026046.1      GI:19684094
KEYWORDS      MGC.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1367)
REFERENCE
Strausberg,R.D., Feilgold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,C.M., Shemen,C.M., Schuler,G.D.,
Altshul,S.F., Zeeberg,B., Buerow,K.H., Scheffer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diachenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Caavaant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshlyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahay,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmitz,J., Myers,R.M.,
Butterfield,Y.S., Kizylinski,M.I., Skalska,U., Smallue,D.B.,
Scherer,A., Schein,J.E., Jones,S.J., and Matra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

```

JOURNAL

PUBMED 12477932
2 (bases 1 to 1367)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (20-MAR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
BC Cancer Agency, Vancouver, BC, Canada
infobgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Lisa Prabhu, Parvaneh Saeedi, JR Santos, Angelique Scherch, Ursula Skalske, Duane Smalins, Jeff Scott, Miranda Tsai, George Yang, Jacques Schein, Asim Siddiqui, Rob Holt, Marco Marra.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAL Plate: 42 Row: f Column: 6
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21389438.
Location/Qualifiers
1. 1367
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:39391 IMAGE:4649261"
/issue_type="Eye, retinoblastoma"
/clone_lib="NIH MGC_16"
/lab_host="DH10B-R"
/note="Vector: pOT87"
1. 1367
/gene="KIS"
/db_xref="LocusID:127933"
159. 1193
/gene="KIS"
/codon_start=1
/product="kinase interacting stathmin"
/protein_id="AAH26046.1"
/db_xref="GI:19684095"
/translacion="MAGSGCAGCABPPRFLFAFRMLQVOSRLGSGSSAVYVRCCG
NPGSPGAKOPLPCTGTGAASAAYGPKERALEOLQGHNTIVLGVPTIHSP
NPSRCLLELDIVSYSELLYSHOGCSKMTQKCARVYLEALILHGGVYHDLK
PRNLTWLAENECFKLIDFGLSFKENQDVYTIQTDYRAVEALQCLQAGLQSDTE
CTSAVDLMSGIIILEMFGMKLKHVRSGEMVNSAYIDHIFAKAVNNAIPAYH
LRLLKSMHLDPSRRIPAEMALCSPPFSIPFAPHIEDVLMPTPLRLNLVDDVYL
ENEEVEGIC"
ORIGIN
Alignment Scores:
Pred. No.: 7.45e-170 Length: 1367
Score: 1990.00 Matches: 385
Best Local Similarity: 92.14% Conservative: 2
Best Local Similarity: 91.67% Mismatches: 2
Query Match: 90.21% Indels: 31
DB: 9 Gaps: 1
US-10-798-532-4 (1-419) x BC026046 (1-1367)
QY 1 MetAAGlySerGlyCysAlaATPGlyAlAGluProProbaRphLeuGluAlaIapheGly 20
DB 159 ATGGCGGATCCGAGCTGGCGCGCGAGCCGCCGCTTTCTGGAGGCTTCGGG 218

QY 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyraArgVal 40
DB 219 CGGCTGTGGCAGATGACAGAGCCGCTCGGGTACCGGCTCTCCGCTCGGTATCGGGTT 278
QY 41 ArgCysCysGlyAsnProGlySerProProGlyAlaAlaArgGlnPheLeuProProGly 60
DB 279 CGCTGCTGGGGAACCTGTGCTGGCCCCCGGCGGCTTCAACAGTTCTTGGCCGAGGA 338
QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrglyPheArgLyGlnArgAlaAlaLeu 80
DB 339 ACCACCGGGGCTGGCGCTCTGCCCGCAGTATGTTTCCGAAAGAGGGCGCGCTG 398
QY 81 GlnGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrglyValPheThrIleHisPhe 100
DB 399 GAACAGTTGACAGGATCAGAAACATCGTGACTTTGTATGAGTGTTCATTCACCTTT 458
QY 101 SerProAsnValProSerArgCysLeuLeuLeuGlnLeuLeuAspValSerAlaGlyGlu 120
DB 459 TCTCCAAATGTGCCATCAGCTGTCTGTGTTGAATCTGAGATCAGTGTTCGGAA 518
QY 121 LeuLeuLeuTyrsSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
DB 519 TTGCTTTTATATTCAGTACAGGGGTTTTCATGTGATATACAGATTGTCCCA 578
QY 141 AspValLeuGlnAlaLeuAlaPheLeuHisHisGlnGlyTyraValHisAlaAspLeuTy 160
DB 579 GATGTTTGGAGCGCTTCTTTCTTCATCATGAGGCTATGTCATCGGACCTCAAA 638
QY 161 ProArgAsnIleLeuTrpSerAlaGluAsnGlyCysPheValLeuIleAspPheGlyLeu 180
DB 639 CCACGTAACATATTTGGAGTGCAGAGATGATGTTTAACTCATTTACTTGGACTT 698
QY 181 SerPheLyGlnGlyAsnGlnAspValLyserTyrlleGlnThrAspGlyTyraArgAlaPro 200
DB 699 AGCTTCAAGAGGACATCAGATGTAAATATTCAGACAGACGGATTCGGCTCCA 758
QY 201 GlnAlaGlnLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGlyCysThr 220
DB 759 GAAGCAGAAATTCAAATTTGCTTGGCCGAGGCTGCTCAAGAGATCAGAAATTAAC 818
QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnMetPheSerGlyMetCys 240
DB 819 TCAGCTGTGATCTGGAGCGCTAGGAAATCATTTTACGGAATGTTTCAGAAATGAA 878
QY 241 LeuLyHisIleThrValArgSerGlnGlyTrpLyAlaAsnSerSerAlaIleIleAspHis 260
DB 879 CTGAACCATACAGTCAATCTCAGAAATGGAAGCAACAGTTTCTACTATTCATCAC 938
QY 261 IlePheAlaSerLyAlaValAlaValAsnAlaAlaIleProAlaTyrlleLeuArgAspLeu 280
DB 939 ATATTGGCAGATTAAGCATGTGTAATCGCGAATTCACGCTTATCACTTAAGACACTT 998
QY 281 IleLySerMetLeuHisAspAspProSerArgAlaIleProAlaGlnMetAlaLeuCys 300
DB 999 ATCAAAAGATCTTCATATATATCCAAACAAAGAAATTCGCTGAATATGCAATTTGTC 1058
QY 301 SerProPhePheSerIleProPheAlaProHisIleGlnAspLeuValMetLeuProThr 320
DB 1059 AGCCCATTTTATAGCATTCCTTTTGGCCCTCATATTGAAGATCTGTCATGCTCCACT 1118
QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrlleGlnAlaGlnGlyGlyTy 340
DB 1119 CCAGTCTTAAGACTGCTGAATGTGCTGATGATGATTAATCTTGAGAAATGAAGCAATAT 1178
QY 341 GlnAspValValGlnAspValLyserGlnGlyCysGlnLyTyrlleProValValSerLeu 360
DB 1178 ----- 1178
QY 361 LeuValProLyGlnAsnProGlyArgGlyGlnValPheValGlyTyraAlaAsnAlaGly 380
DB 1179 -----GAAAGTCTTTGTGAGTATGCAAAATGCTGG 1208

QY 380 YAPSerLYsAlaAlaGlnLYsLeuThrGlyArgMetPheAspGlyLYsPheValVa 400
Db 1209 TGATTCCAAGCTCGCGAAGAAATTACTGATCGAAGATGTTGATGGAACTTTGTGT 1268
QY 400 lAlaThrPheTYrProLeuSerAlaTYrLYsArgGlyTYrLeuTYrGlnThrLeuLeu 419
Db 1269 GCGTACATTTCTACCCGCTGAGTGCTCAAGAGGGGATCTGTATCAACTTTGCTT 1326

RESULT 15
HSM805327 2940 bp mRNA linear PRI 12-JUL-2002
LOCUS Homo sapiens mRNA; cDNA DKFZp6710511 (from clone DKFZp6710511).
DEFINITION AL834136
ACCESSION AL834136 GI:21739600
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2940)
AUTHORS Koehler, K., Beyer, A., Mewes, H.W., Weil, B. and Wiemann, S.
JOURNAL Direct Submission
COMMENT Submitted (09-JUL-2002) 1, D-85764 Neuherberg, GERMANY
clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Berlin-Charlottenburg, GERMANY; Email: clone@zrp.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/
location/Qualifiers
FEATURES
source 1..2940
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp6710511"
/db_xref="taxon:9606"
/map="1"
/clone="DKFZp6710511"
/issue_type="lymph node"
/clone_id="667 (synonym: hlmo2). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
2923

polyA_site
ORIGIN
Alignment Scores:
Pred. No.: 2,09e-160 Length: 2940
Score: 1880.50 Matches: 388
Percent Similarity: 63.46% Conservative: 1
Best Local Similarity: 63.30% Mismatches: 2
Query Match: 85.70% Indels: 223
DB: Gaps: 1
US-10-798-532-4 (1-419) x HSM805327 (1-2940)

QY 29 LeuGlySerGlySerSerAlaSerValTYrArgValArgCysCysGlyAsnProGlySer 48
Db 1 CTGGGTAAGCGGCTCTCCGCTCGCTGTATCGGGGTTGCTGCGGCAACCTGGGCTCG 60
QY 49 ProProGlyValAlaLeuArgGlnPheLeuProProGlyTYrThrGlyValAlaAlaSerAla 68
Db 61 CCCCCCGGCGCCCTCAAGCAGCTTCTGGCCCGCAAGAACACCGGGGCTGGGCTCTGCC 120
QY 69 AlaGlyTYrGlyPheArgLYsGlyArgAlaAlaLeuGlnGlnGlyAlaSerAsn 88
Db 121 GCGCAGATGTTTCCGCAAGAGAGGGCGCGCTGGAACAATTGCGAGGTCAAGAAC 180
QY 89 lIleValThrLeuTYrGlyValPheThrIleHisPheSerProAsnValProSerArgCys 108
Db 181 ATCGTACCTTGTATGAGAGTGTATTAACAATCACTTTTCCAAATGTGCATCAGCTGT 240

QY 109 lIleLeuLeuGlnLeuLeuAspValSerValSerGlnLeuLeuTYrSerSerHisGln 128
Db 241 CTGTTCCTGAATCTCGAGATGTCAGTGTTCGAAATGCTCTTATATTCACATCCAG 300
QY 129 GlyCysSerMetTrpMetIleGlnHisCysAlaArgAspValLeuGlnAlaValPhe 148
Db 301 GGTGTTCCATGTGATGATGATACAGCATGTGCGCGAGATGTTTTGAGAGCCCTGCTTTT 360
QY 149 lIleHisGlnGlyTYrValHisAlaAspLeuLYsProArgAsnIleLeuTrpSerAla 168
Db 361 CTTCATCAATGAGGCTATCTCCATGCGGACCTCAACACATTAATTTGTGAGATGCA 420
QY 169 GluAsnGlnCysPheLYsLeuIleAspPheGlyLeuSerPheLYsGlnGln---- 187
Db 421 GAGATGAATGTTTTAACTCATTCGCTTGGACTTGAAGAGGCATATCA-GGT 479
QY 187 ----- 187
Db 480 AAGAATAAAGCTTTCTTCTCTGCACTTAAATGATGTGTTGAAGCCAAATATTT 539
QY 187 ----- 187
Db 540 TTAATGACAAGGATGTTAGTAGGGGCTTACAACTCATTTTAACTTAATTAAGTAA 599
QY 187 ----- 187
Db 600 TGCCAGGGATGCAGTAAATCAAAAGATTTCAGTCCACCGGTAAAAAAGACATTA 659
QY 187 ----- 187
Db 660 TGTTGGGTGAGTAGGAGAGACGATGAAGCATGAGCTGTGAATTTATTAAGTAC 719
QY 187 ----- 187
Db 720 TTAACCTAGATGAATGTTTTTTTAAATTTATGCGTGTGGCTAATGCTGAGTTACT 779
QY 187 ----- 187
Db 780 ATTCTGTGCTCTTCACTGCTGTGGGCTGTATCTCGTGCAAGATGATATGATGTTCTG 839
QY 187 ----- 187
Db 840 CCAATCCTCTTTTCCATGATGTTATCATTTCTTTCAGATGCTTATTAACAGAGTG 899
QY 187 ----- 187
Db 900 CCTAATAGGTAAAGACAATCCAAAGAGACTTCTTATGAGCGCTATATGAAATGTAC 959
QY 187 ----- 187
Db 960 AGCTTCCTCTGAGAACTGCCCAAGTTAATGAACAAAGAAATTTATGTAATCAAGACTAA 1019
QY 187 ----- 187
Db 1020 AAATTTTACATTAGCTTGGAGATGGGTTTACTGCACTTTAGGGGAAGTAATTCAC 1079
QY 187 ----- 187
Db 1080 GTAGAGATGATAGAAAGGAGAGCTTTTATTTGTTGTAATATTACTCATCTTTT 1139
QY 188 ---- AspValLYsTYrIleGlnThrAspGlyTYrArgAlaProGlnAlaGlnLeuGlnAsn 206
Db 1140 TAGAGATGAATGATATTTCAAGACAGCGGATTCGGGCTCCAGAACCAATTCGCAAAAT 1199
QY 207 CysLeuAlaGlnAlaGlyLeuGlnSerAspThrGlnCysThrSerAlaValAspLeuTrp 226
Db 1200 TGCTTGCCCAAGGCTGCTGCAAGTGTATCAGATGATACCTCAGCTGTGATCTGTGG 1259
QY 227 SerLeuGlyIleIleLeuLeuGlnLecPheSerGlyMetLYsLeuLYsHisThrValArg 246
Db 1260 AGCTTAGAATCATTTTACTGGAATATGTTCTCAGGAATGAACTGAACATACATCAGTCAGA 1319


```
QY 247 SerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHisIlePheAlaSerLysAla 266
Db 1320 TCTCAGGAATGAGAGCAAAACAGTTCTGCTATTATTGATCACAATATTGCCAGTAAAGCA 1379
QY 267 ValValAsnAlaAlaIleProAlaIleTyrHisLeuArgAspLeuIleLysSerMetLeuHis 286
Db 1380 GTGGTAAATGCCGCAATTCCAGCCTATCACCTTAAGAGACCTTATCAAAAGCATGCTTCAT 1439
QY 287 AspAspProSerArgArgIleProAlaGluMetAlaLeuCysSerProPhePheSerIle 306
Db 1440 GATGATCCAGACGAAAGAAATTCCTGCTGAATGCAATGCAATGCAATGCAATGCAATG 1499
QY 307 ProPheAlaProHisIleGluAspLeuValMetLeuProThrProValLeuArgLeuLeu 326
Db 1500 CTTTGGCCCTCATATTGAAGATCTGTCATGCTCCCACTCCAGTGTAGAGACTGCTG 1559
QY 327 AsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyrGluAspValValGluAsp 346
Db 1560 AATGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1619
QY 347 ValLysGluGluCysGlnLysTyrGlyProValValSerLeuLeuValProLysGluAsn 366
Db 1620 GTAAAGAGAGAGTGTCAAATAATGACCACTGCTATCTTACTTGTCCAAAGCAAAAT 1679
QY 367 ProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGlyAspSerLysAlaIleGln 386
Db 1680 CCTGCCAGAGACAGTCTTGTGTTGATGATGCAAAATGCTGTGATTCCAAAGCTGCGCAG 1739
QY 387 LysLeuLeuThrGlyArgMetPheAspGlyLysPheValValAlaThrPheTyrProLeu 406
Db 1740 AATTTACTGACTGGAAGAGATGTTGATGGGAGAGTTGTTGGCTACATTCTACCCGCTG 1799
QY 407 SerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
Db 1800 AGTGCTACAGAGCGGATATCTGTATCAAACTTGCTT 1838
```

Search completed: November 29, 2004, 18:11:29
Job time : 4618 secs